Epidemiology and molecular diversity of influenza A viruses in pigs

A DISSERTATION
SUBMITTED TO THE FACULTY OF GRADUATE SCHOOL
UNIVERSITY OF MINNESOTA
BY

Carlos Andres Diaz Jimenez

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Dr. Montserrat Torremorell.

November 2015

Acknowledgements

In August 2010, I started my graduate studies at the University of Minnesota, and today, I lack enough words to thank everybody who help me achieve my Ph.D. I would like to start thanking my parents Odila and Milciades. Thank you for your unconditional support and love throughout my life. Thanks for my education and values. Thanks for teaching me how to pursue my dreams and work hard to achieve them. I am proud of the man I am today and that is because of you. I have enjoyed this journey and there is no way I would have accomplished my thesis work without your support and encouragement. Furthermore, I would like to thank my sister Claudia for her love and friendship. Thanks for showing me that when things get hard you try harder, I have learned so much from you. Also, thanks to all the members of my extended family that have been in touch through my journey; I appreciate it.

Thank you to my advisor, Dr. Montserrat Torremorell for the opportunity to become part of her research group. Thank you for your trust, patience, believing in my ideas, and helping me ground, when necessary. I will always be grateful to you. I have grown so much as a swine veterinarian, a researcher, and more importantly as a person under your guidance. Additionally, thanks to my graduate committee because you made me think beyond my limits. Thanks for sharing your knowledge and experience with me, and for your invaluable contributions to my thesis work. Thank you, Dr. Srinand Sreevatsan for the opportunity to discuss in depth my findings and make me think further to understand them. Thank you, Dr. Marie Culhane for sharing all I wanted to know regarding influenza A virus and swine diseases. Thank you, Dr. Peter Davies and Dr. Claudia Muñoz because you made me find my research fascinating since it was driven by the epidemiology of infectious diseases. Your knowledge and experience was very appreciated. Furthermore, I would also like to acknowledge all other faculty and staff at the University of Minnesota that contributed to my research and success,

especially Dr. Mark Rutherford, Dr. William Hueston, Lisa Hubinger and Katherine Barry. Special thanks also to Dr. Dario Mogollón who graduated from the University of Minnesota and has been my career mentor for almost 10 years.

My life experience at the University of Minnesota and the Twin cities would have not been as fun without my friends. Thanks to my Colombian friends who gave me their support from the distance especially to Juliana, Violeta and Juan. Thanks Juli for visiting, especially in the winter. That is love when you are not from Minnesota. Furthermore, thanks to my friends from all over the world. You have been an incredible part of this mind expansion. Thank you, Michael for your unconditional support and help, I am grateful for sharing so many different life experiences. Special thanks also to Ana, Doug, Carmen, César, Anna, Maria José, Devi, and My for your friendship and support, it has been so much fun. Additionally, thanks to all graduate students who made this journey fun and entertaining, specially Tiffany, Matt, Nubia, Steve, and Niti. Thanks also to all the students in the swine group and visiting scholars for your help in my projects.

Finally, I would like to acknowledge all my funding sources. First, I would like to thank the Colombian Government and COLCIENCAS for the Caldas fellowship. Thanks to the Swine Disease Eradication Center for the Carlos Pijoan Scholarship and the MNDrive initiative for the MNDrive fellowship. Thanks to the Graduate School, University of Minnesota for the Doctoral Dissertation fellowship and the College of Veterinary Medicine, University of Minnesota for the travel awards granted to present my results nationally and internationally. Thanks to the Centers of Excellence for Influenza Research and Surveillance (CEIRS), especially the Minnesota CEIRS (MCEIRS), and National Pork Board for the funding, and tanks to Pipestone Veterinary Clinic and Swine Veterinary Center for their collaboration in this thesis projects, especially to Dr. Gordon Spronk, Dr. Spencer Wayne, and Dr. Paul Yeske. Last but not least, thanks to all the pork producers and farm staff that made this research possible.

Dedication

To my parents Odila and Milciades for their unconditional love and support.

Table of Contents

List of Tables

Chapter 5:

Chapter 6:

List of Figures

Chapter 5:

Chapter 6:

General introduction

Influenza A viruses (IAVs) infect many animal species including human (1), avian (2), swine (3), equine (4), canine (5), feline (6), some marine mammals (7) and bats (8). Zoonotic IAV infections happen worldwide and are considered a major public health risk because IAVs can cause human pandemics (9). A novel IAV caused the 1st human pandemic of the 21st century and highlighted the role of the pig in the ecology of IAVs because the 2009 pandemic virus emerged from IAVs circulating in pigs in North America and Asia (10). In humans and pigs, IAVs are a main cause of respiratory disease and multiple interspecies transmission (human-pig) events have been documented (11-13). However, only one swine-origin influenza virus, the 2009 pandemic virus, has adapted to humans and acquired human-to-human transmission (9, 10, 14). In contrast, several human influenza viruses have become established in swine populations (15). More extensive studies of the molecular biology and epidemiology of IAVs in pigs could help determine the emergence of novel IAVs with zoonotic and pandemic potential.

Different epidemiological studies have evaluated the exposure, transmission, and distribution of swine IAVs around the world using serological tests or more recent IAV detection methods (3, 16-18). However, there is a gap in knowledge about what happens at the herd level, specifically between the epidemiological findings during IAV infection of pigs and the molecular information obtained from the virus. Swine IAVs are distributed worldwide and direct contact with infected pigs is considered the main route of transmission (3) although fomites (19) and airborne transmission may also play a key role in the spread of swine IAVs (20). Additionally, the main antigens of IAVs (hemagglutinin and neuraminidase) may differ by more than 60% (21, 22) illustrating the wide diversity of the viruses. Furthermore, in the last 20 years the genetic makeup of swine IAVs has changed significantly (17, 23). Nevertheless, viruses cannot replicate without their host and infected individuals can transport the virus to distant locations. Therefore, pig movements (regionally and globally) are associated with the genetic diversity and evolution of swine IAVs (13, 24) and highlight the significance of the host on virus

evolution and diversity. Moreover, pig production systems have evolved into highly efficient farms that are concentrated in certain parts of the world (25-28). Pig turnover rates and animal movements within and between farms make viral infections harder to control. Although animal flows and biosecurity practices are in place to prevent other viral infections (29-32), the factors that increase the risk of IAV infection are not clearly understood.

IAVs are commonly found among pigs (17) and represent a significant cost to producers (33, 34). In addition, IAVs can persist at the herd level for prolonged periods of time and one or more IAV subtypes can circulate over time within the same farm (3, 17, 18). However the epidemiological characteristics and molecular traits of IAVs that allow the long-term persistence of the virus in pig farms are not clearly understood. It is not clear if the genetic evolution within farms is responsible for virus persistence at the herd level or if there are repeated introductions of IAVs into pig herds. Moreover, different pig subpopulations are housed within the same farm (e.g piglets, sows, and replacement animals) and it is not clear if these subpopulations harbor the same or different IAVs over time. Furthermore, in breeding herds replacement animals (gilts) are introduced on a regular basis but their role in IAV epidemiology is unknown. Gilts and newborn piglets could represent important subpopulations in which IAVs may replicate continuously over time.

Once pigs are born they may be exposed to different IAVs under a variety of immunological circumstances and develop diverse immune responses accordingly. Pigs are born naïve to all IAVs because infection does not take place in-utero (35). Furthermore, there is no antibody transfer before birth (36). Then, after birth maternally derived antibodies (MDA) can be transferred to suckling piglets in colostrum, depending on previous exposure of sows to IAVs. The effect of MDA on IAV genetic diversity is not clearly understood. However, the genetic diversity of IAVs during infection of pigs with active immunity to multiple IAVs has not been characterized.

Understanding the transmission dynamics and genetic diversity of swine IAVs at the herd level should allow us to design better health interventions to control the disease in pigs, minimize IAV's effect on swine health and production, and reduce the public health risk. Moreover, understanding the molecular dynamics of IAV genome at the individual and population levels during natural and experimental infections will help us understand the genetic diversity among individual hosts.

Overall we hypothesize that the immune status of the pig is a key driver of the molecular evolution of the virus and that viral diversity alters the course of IAV infection. We also propose that different pig subpopulations on swine farms play unique roles in the persistence of IAVs over time. Therefore the main objective of this dissertation was to characterize the genetic diversity of IAVs during infection of pigs to better understand the epidemiology of swine influenza, focusing in particular on what happens at the herd level. Classical epidemiological methods and novel experimental designs were integrated with deep genome sequencing technologies and bioinformatics algorithms to produce robust evidence that supports the genetic plasticity of IAVs and contributes to the understanding of virus persistence at the herd level. The specific aims of this dissertation are to:

1. Define patterns of IAV infection in pig subpopulations in IAV infected breeding herds.
2. Assess and compare the genetic diversity of IAVs isolated over time from endemically infected breeding herds to estimate how long IAVs can persist at the population level.
3. Evaluate the antigenic drift of an H1N1 IAV during infection of weaned pigs with or without passive immunity.
4. Evaluate the genetic diversity of the complete genome of a triple reassortant H1N1 IAV population during experimental infection of vaccinated pigs.

5. Characterize, under field conditions, the epidemiology and molecular traits of IAVs infection in pigs after weaning.

These objectives were achieved in different studies that focused first on commercial pig-breeding herds and then in weaned pigs under both experimental and field conditions. Understanding the epidemiology (objective 1) and genetic diversity (objective 2) of swine IAVs at the breeding herd level is crucial to minimize the distribution of IAVs to other locations after pigs are weaned. Then, studies on weaned pigs were conducted to understand the genetic variability of IAVs under different immunological conditions. These studies included determining the genetic diversity of the hemagglutinin of an H1N1 IAV was studied during infection of pigs with or without maternal immunity (objective 3); then, the complete genome of the virus was studied during IAV infection of pigs with immunity to different IAVs (objective 4); and lastly we studied the epidemiology and molecular traits of different IAVs during infection of pigs after weaning under field conditions (objective 5).

In these studies, models of IAV transmission and evolution included immune and non-immune pigs, and more importantly, pigs with passive immunity. Passive immunity in pigs plays a central role in IAV control and transmission since pigs with passive immunity sustain virus replication despite the lack of clinical signs (37, 38), which may represent ideal conditions for virus evolution. Moreover, two longitudinal field studies were developed in order to fully represent complex scenarios in swine IAV ecology. The combination of experimental and field settings to address complex questions using novel technologies yielded robust evidence regarding swine IAV epidemiology and evolution, which resulted in the translational application of the research findings to the swine industry.

Chapter 1: Literature review

Sections of this chapter have been published in:

Torremorell, M., Allerson, M., Corzo, C., Diaz, A., Gramer, M., 2012. Transmission of Influenza A Virus in Pigs. Transbound Emerg. Dis. 59 (Suppl. 1) (2012) 1–17.

Influenza A virus (IAV) and susceptible host species:

Influenza A viruses (IAVs) are enveloped single stranded negative sense RNA viruses (-ssRNAv) that belong to the *Orthomyxoviridae* family (39). IAV genome is segmented in eight gene segments namely polymerase B2 (PB2, segment 1), polymerase B1 (PB1, segment 2), polymerase A (PA, segment 3), hemagglutinin (HA, segment 4), nucleoprotein (NP, segment 5), neuraminidase (NA, segment 6), matrix (M, segment 7), and non-structural protein (NS, segment 8). From these gene segments, at least eleven different proteins are translated (Table 1) and among them the hemagglutinin (HA) and neuraminidase (NA) proteins are considered the main antigens of the virus. HA and NA also used to classify the IAV subtypes and to guide IAV vaccine selection in many animal species including humans (40). There are at least 17 different HAs and 9 NAs (8, 9) although only a few HA-NA combinations have been reported in pigs (16, 18). Overall, the genetic diversity and evolution of IAVs is determined by multiple factors including the molecular characteristics of the virus, the host species, the immune responses, epidemics, and pandemics (1, 2, 9, 10, 41, 42).

Avian species within the orders *Anseriformes* and *Charadriformes* are the natural reservoirs for IAVs (9). However, IAVs can infect other animal species including human (43), swine (3), equine (4), canine (5), feline (6), some marine mammals (7) and bats (8). The molecular mechanisms that allow IAV to cross inter-species barriers are not clearly understood (9, 44). However, the HA affinity for cell receptors plays a key role in host species range (39, 45-47). While avian viruses have greater affinity to bind to cell receptors that have sialic acid linked to the penultimate galactose in an α2-3 configuration (expressed more frequently in birds), human IAVs bind to α2-6 configured receptors (expressed more frequently in mammals). The HA is not the only determinant for IAVs host species ranges. For example, polymerases from avian IAVs do not perform efficiently in mammals (48). Moreover, pigs have both types of receptors for IAVs (α2-3 and α2-6) hence they are susceptible to infection to some mammalian and avian IAVs

(49). Interestingly, humans also possess both receptors on their cells for avian, human, and swine IAVs and zoonotic infections with avian or swine IAVs are continuously reported around the world (11, 50-52).

Zoonotic IAVs are a major public health risk because such IAVs can lead to pandemics (10, 53). All IAV pandemics after the Spanish flu emerged from zoonotic and reassorted IAVs (9). Whether the 2009 pandemic strain emerged from a reassortment in pigs, humans or other animal species is still not clear. However, this pandemic virus highlighted the significance of swine IAVs for public health. Nevertheless, the 2009 pandemic virus has been the only swine IAV able to acquire human-to-human transmission after zoonotic infection (9, 10). In contrast several human IAVs are well established in swine populations (13, 15) and have contributed to the current diverse genetic landscape of swine IAVs (54-56).

Influenza A virus evolution and immune selection:

Overall, RNA viruses change rapidly over time and IAVs are not the exception (44). A non-proof reading RNA-polymerase allows nucleotide mutations to accumulate over time during viral replication (48). The divergence of HA and NA genes over time is known as antigenic drift (39). In humans, HA has shown different rates of antigenic drift between subtypes, with H3 changing faster than H1 variants (41). Moreover, substitution rates of human H3, but not H1, are characterized by long periods with higher numbers of synonymous mutations (stasis periods) followed by shorter intervals of rapid evolution where new dominant variants are established (57). Additionally, human H3 viruses have gene segments with different evolutionary patterns and origins that persist due to migration of meta-populations (58). Moreover, one study showed that IAV gene segments could evolve at different rates in different host species (59), which might be associated with the immunological responses to IAVs. Some results

indicate that human IAVs have higher mutation rates than swine IAVs (60) while a recent study indicated the opposite (61).

The immune response to IAV is associated with antigenic drift in several animal species (41, 62, 63). Immune responses mediate IAV divergence over time because certain IAV strains (antigenically divergent) can evade the immune response and give origin to new genetic lineages. In swine, homologous immunity can minimize or prevent IAV transmission (38, 64). However, the effect of heterologous immunity on transmission and replication is not clearly understood. In animal models other than pigs, the original antigenic sin (OAS) phenomenon indicates that the immune response to new IAVs could be impaired if a previous exposure to a different IAV had happened (65). To my knowledge the OAS has not been investigated in detail for IAVs in pigs. Only one study indicates that the OAS phenomenon is not associated with the vaccine-associated enhanced respiratory disease (VAERD) observed in pigs during some IAV infections (66). Studies of OAS in rabbits with antigens differing more than 33 – 42% are not able to efficiently stimulate the memory immune system, and indicate that the homology between antigenic proteins is associated with the extent of immune recall (67). In humans, masking of epitopes of new IAV antigens with both pre-existing antibodies and with antibodies developed during the immune response drive the extent of the humoral immune response to IAVs (68). Furthermore, the memory humoral response to IAVs will depend on the amount of free antigen stimulating the immune system. However, IAV infection appears to be better at inducing antibodies to previous IAV antigens than IAV vaccination. Nevertheless, the characterization of the plasmoblast antibody response to IAVs in humans showed that vaccination to IAVs could induce the production of antibodies against IAVs circulating in previous seasons that were not contained in the vaccine used (69). These results in other animal species suggest that in pigs the interaction between IAVs and the immune status of the population could also play a key role in the emergence, persistence, and subsidence of IAVs over time. However, the antibody repertoire to IAVs in pig

populations and its relationship with IAV evolution is poorly understood and should be further investigated.

The segmented genome of IAV allows gene reassortment to happen when two or more IAV particles infect the same cell (9, 39). Genetic reassortment can change abruptly the antigenic properties of IAVs (antigenic shift). Furthermore, reassortant IAVs are of public health concern because reassorted IAVs can cause pandemic IAV infections (9, 10). Genetic reassortment can also change viral pathogenesis without exchanging antigenic genes (when only internal genes are swapped) and complicates even more the understanding of IAV evolution and diversity within and between different species. Even though gene segment exchange has been demonstrated within and between species, little is known on the IAV gene flow in populations of endemically infected pigs. In other species reassortment appears to be random when IAVs are closely related to each other (70) and restrictive (not-random) when viruses from different genetic lineages exchange gene segments (71, 72). Nevertheless, the molecular mechanisms that drive swine IAV evolution and diversity could determine the health intervention practices required to minimize the impact of the disease in swine health and production and to reduce the public health risk.

Epidemiology and genetic diversity of swine IAVs:

Swine influenza was first described in 1918 coincidently with the Spanish flu pandemic (16). However, IAVs were not isolated from pigs until 1930 (classical H1N1, cH1N1). To date multiple IAV subtypes have been recovered from pigs (H1N1, H3N2, H1N2, H3N1, H4N6, H7N2, H9N2, and H5N1(3)). However only a few of these different subtypes transmit efficiently between pigs and are considered endemic in swine populations (H1N1, H3N2, and H1N2). Multiple IAV alleles (sequence variants of the same virus) can co-exist during IAV infection of pigs (73) and the same virus can evolve differently in the upper and lower respiratory tract of pigs (74).

The most prevalent swine IAV subtypes in North America (H1N1, H1N2, and H3N2) have phylogenetic origins in the classical H1N1 (cH1N1), the triple H3N2 reassortant, the human H3N2 and the Eurasian swine H1N1 IAVs (13, 54-56, 61, 75, 76). After the first isolation of IAV in 1930, and for almost 70 years, almost all swine IAVs isolated in North America were related to cH1N1. Then, in the late 20th century multiple avian and human IAVs became established in swine populations (16, 61). However, it is not clear why "foreigner" IAVs (from humans and birds) only became established in swine populations until the last quarter of the last century and not before. Nevertheless, serological evidence indicates that human H3 were circulating within swine populations before becoming established (16).

These multiple introductions of avian and human IAVs changed the genetic landscape of swine IAVs after 1998, when an outbreak of swine IAV in North Carolina and in Minnesota reported the introduction and maintenance of at least two different genotypes of H3 subtype into the US swine population. After this outbreak several different subtypes (H3N2, H1N2, human H1N1, and H3N1) containing triple reassortant internal genes (TRIG) cassette were identified. This TRIG cassette was composed of PA and PB2 genes from the avian IAVs, NS, NP, and M from the cH1N1, and PB1 gene from the human IAVs (16). Interestingly, this internal gene constellation was maintained over time within many swine IAVs regardless of their subtype. However, after multiple reassortment events novel IAV gene constellations have emerged in North America (54, 75, 76). The frequency of reassortment among swine IAVs is not clearly understood. Only one study in Europe demonstrated that the rate of IAV reassortment is not the same for different IAV subtypes in swine but further investigation is required on this topic (72). Moreover, in 2009 the pandemic IAV placed "the pig" at the center of the "the mixing vessel" discussion because the pandemic virus contained gene segments from IAVs circulating in pigs in Asia and North America (10). Furthermore, during the pandemic more reassortment events happened with swine IAVs and illustrated how commonly virus

reassortment could take place among current and newly introduced viruses. Nevertheless, after the 2009 pandemic not all genes from the pandemic IAV remained circulating at the same level in the U.S pig population (75).

At the HA level, swine IAVs in North America cluster in six antigenically and phylogenetically distinct H1 clades (α, β, γ1, γ2, δ1, and δ2 (56, 76)) and four H3 clusters (I, II, III and IV (54, 75)). The genetic diversity of swine IAVs in the USA has been associated with the geographical distribution and movement of pigs (13, 24). Human viruses have become established in swine populations (15) and the Midwest appears to serve as an ecological niche for swine influenza viruses to reassort after they have originated from different areas of the country (13). Moreover, international trade of pigs is also considered a predictor of IAV flow from the USA and Europe to Asia (24, 77). IAV evolution is marked by the interaction of rapid mutation rates, viral selection, reassortment and worldwide epidemiological factors (9). In pigs, however, the host or herd factors influencing IAV evolution and diversity are not completely understood.

Moreover, weaned pigs are main sources of IAVs and as a large subpopulation in the herd they appear to be an important herd factor to study. Weaned pigs can play a key role in the persistence of IAVs over time, They can be infected without showing any clinical signs (37), and transporting the virus to multiple geographical locations after weaning. Additionally, piglets with or without passive immunity can shed IAVs (38) and multiple IAV subtypes can circulate in piglets before and after weaning (17, 18), which may facilitate the emergence of reassortant IAVs. However, little is known about genetic diversity and selective evolution of IAV in weaned piglets with or without immunity to IAVs. To my knowledge, there is only one published study that evaluated the effect of vaccination on the genetic evolution of IAV in pigs and in that study no differences were found in the mutation rate between pigs with or without active immunity to IAVs (73). However, the unknown variety of immune memory to IAVs

among pig populations could be associated with IAV genetic diversity and distribution.

Swine IAVs are distributed worldwide and in the Midwestern USA 90% of the farms with growing pigs are positive to IAVs (17). In North America swine influenza was considered an epidemic disease with higher incidence rates in the late fall and early winter (16). However, the seasonality of swine influenza is not as marked as human influenza infections in temperate regions. Recent studies indicate that IAVs could be recovered from swine year round in endemically infected herds (17, 78, 79). Moreover, IAV infections in swine have a short incubation period (1-3 days) and animals recover within one-week post infection (16, 80, 81). Infected pigs are either sub-clinically infected or develop non-specific clinical signs including high fever, coughing, sneezing, nasal discharge, decreased food intake, and conjunctivitis. After infection, IAV in pigs results in high morbidity (near 100%) and low mortality (<1%) (3, 16) and its incidence could be associated with the population immune status because transmission rates are different between naïve and immune animals (19, 64).

The main transmission route of IAVs between pigs is direct contact (3). However, aerosols and fomites have been shown to be possible routes of transmission (19, 20). Introduction of infected animals, and pig movement within and between herds can increase the risk of infection in swine populations. Weaned pigs are considered a source of IAVs to other swine populations (37). Additionally, newborn piglets and other naïve populations may play a key role in the reservoir and transmission of IAV within and between pig farms. However, viral persistence is not clear in endemically infected populations.

Prevention and control

IAV prevention and control in pigs is based on biosecurity practices and vaccination. However, strain selection for vaccines is complicated due to the

wide diversity of swine IAVs. Multiple commercial and autogenous vaccines are commonly used and have proven to reduce lesions, clinical signs and virus shedding (21, 64, 82, 83). Transmission might be reduced after vaccination but not eliminated, (64) and will depend on strain relatedness between circulating viruses and those used in the vaccines (21, 82). Furthermore, swine IAV vaccines are not updated systematically based on epidemiological surveillance as they are in humans. Regional dissemination of swine IAVs is highly suspected but its mechanism is still not well characterized. Measures toward the reduction of airborne pathogen introduction should be taken into account, as well as having bird-proofed facilities to keep birds out of pig herds. Furthermore, to better control IAVs, increased surveillance on swine IAVs is required in addition to new policies on IAV vaccination and control.

Swine IAV diagnosis:

Swine IAV diagnosis is basically based on pathology, serology, virus isolation and molecular detection of IAV gene segments (16, 83-85). IAV can induce respiratory clinical signs and lesions in pigs and, although the macroscopic and microscopic lesions associated to IAV are not always specific, they are sometimes highly characteristic. To be prudent, other etiological agents should be listed as differential diagnosis when IAV infections are considered. Furthermore, IAVs are part of the swine respiratory disease complex (86) thus IAVs are usually found with other respiratory pathogens during a respiratory outbreak.

Serological tests are used to measure previous exposure to IAVs and include enzyme-linked immunosorbent assay (ELISA), hemagglutinin inhibition test (HI), and serum neutralizing (SN) tests (87). There are ELISA tests to detect antibodies against IAVs (88, 89) or IAV antigens; both ELISA tests (antibodies and antigen detection) use the NP of the virus because this protein is highly conserved among all IAVs. Additionally, the HA of IAVs agglutinate red blood

cells therefore serum samples could be tested for specific anti-HA antibodies using an HI test (90, 91). HI tests are useful for vaccine selection and are used by the World Health Organization to estimate when human IAV vaccines should be updated (40). SN tests measure the capability of a serum sample to inhibit IAV infection in cell cultures where the virus could grow in the absence of specific antibodies.

Virus isolation is considered the gold standard for IAV diagnosis. However, molecular detection of IAVs can be more sensitive for IAV detection. Swine IAVs are usually grown in Madin-Darby canine kidney cells where cytopathic effect (CPE) can be observed within a week of cell culture infection (92). ELISA tests or molecular detection methods are used to confirm IAV isolation after CPE has been observed. Moreover reverse transcription real time polymerase chain reaction (RRT-PCR) is routinely used in the USA for the detection of swine IAVs. The United States Department of Agriculture (USDA) official test to detect swine IAVs by RRT-PCR targets the matrix gene of classical swine and pandemic IAVs (84). Furthermore, subtyping PCRs are available to differentiate H1N1 from, H1N2, H3N1 and H3N2 viruses. Nevertheless not all IAV positive samples can be subtyped using these methods, demonstrating the diversity among swine IAVs. IAV subtyping is inadequate for understanding the diversity and evolution of IAV within and between pig populations because viruses clustering within the same subtype could have evolved from different IAV lineages and could have different antigenic properties. Therefore, sequencing of swine IAV HA, NA and M gene segments has increased considerably in the last decade, especially after the 2009 IAV pandemic.

Next generation sequencing technologies and complete IAV genome sequencing:

Each IAV gene segment can be amplified in RT-PCR reactions and the PCR

amplicons can be used for genetic sequencing (85, 93). Different sequencing technologies are available for IAV sequencing. In the past, IAV gene segments were mostly amplified using “walking primers” and then sequenced using Sanger sequencing technologies. The Sanger technology allows long sequencing reads but low coverage (e.g 1 to 10X) because of the time required to read each base pair during the sequencing process (94, 95). Gene assembling using Sanger is easier and straightforward because the read length is longer. However, the detection of viral diversity is less likely when polymorphisms are present at low prevalence within the viral population.

In contrast next generation sequencing technologies (NGS) do not require specific primers for genetic sequencing (96-98) and yield extremely high coverage (thousands to millions), where coverage refers to the number of times a given nucleotide position becomes sequenced. The coverage of NGS technologies allows a better characterization of IAVs and the identification of nucleotide polymorphisms of lower prevalence. NGS technologies are able to detect virus variants that are present at 0.1% within the virus population (99). However sequencing lengths are shorter and sequencing errors are higher compared to Sanger sequencing. Therefore sequencing assembling and metagenomes reconstructions are challenging. Nevertheless, new NGS platforms are improving their sequencing read length and new algorithms are available for sequencing assembly (97, 100, 101). Quality control and quality assurance can employ fastqc (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) during the bioinformatics analysis of NGS. During NGS, errors are more likely found within homopolymeric nucleotide sequences (repeats of the same nucleotide) and at the end of longer sequencing reads (98) . Therefore, to avoid overestimation of IAV diversity strict quality control and quality assurance procedures are required to trim sequencing reads with lower sequencing scores, and verify the accuracy of polymorphisms within homopolymeric regions (102).

Moreover, the conserved and complementary 3’ and 5’ ends of all IAV gene segments allow us to amplify the complete genome of the virus in a single RT-PCR reaction (103). IAVs contain eight different gene segments that are very diverse between IAVs. However, all IAV gene segments (including those from different animal species) have 12 to13 nucleotides that are conserved and complementary at the 3’ and 5’ ends of each gene segment. Hence these conserved sequences can be used as primer targets to amplify the complete genome of IAVs in a single RRT-PCR reaction. Since NGS do not require specific primers for gene sequencing, the complete genome obtained from this single RRT-PCR reaction can be sequenced at once and then sorted bioinformatically by gene segment.

Illumina Sequencing (one type of NGS) can yield a large amount (millions) of sequencing reads that range between 50 and 250 base pairs in length (96, 104). These reads can be analyzed as single reads or as pair ends which increases the accuracy of sequencing because takes into account nucleotides found in both directions (forward and reverse) (105, 106). There are two methods to assemble NGS sequencing reads, template based and de-novo assembly (105). In the template-based method, a known sequence template is used to map all NGS sequencing reads obtained. This means that each read is aligned to a template and if a threshold of percent pair wise identity is reached then the sequence is mapped to that region. The consensus (contig) sequence from all reads mapped is extracted and all reads that do not map to the template are discharged. In the case of IAVs one template can be used for each gene segment and one contig can be obtained for each sequence template used. For highly divergent genes (HA and NA) more than one reference template are used to increase the specificity of sequencing mapping. Reference based assembly is faster than de-novo assembling because each read is only compared to the template and mapped or discharged. Additionally, if multiple IAV gene segments are co-circulating within a sample and only one template for that gene segment is used only the most prevalent will be detected (the consensus). If the reference

template used for template based methods is divergent from the sample sequence some genetic variants might be missed. In contrast during de-novo assembly all NGS sequencing reads are compared among them and then all possible consensus (contigs) found are estimated. This process takes longer time because the number of comparisons increases exponentially with the number of NGS reads obtained. De-novo assembling can be useful to discover new IAV lineages.

Swine influenza and the contemporary swine industry:

Although the molecular characteristics of the virus and its natural relationship with the pig can drive IAV evolution it is possible that the structure of the contemporary swine industry is associated with the current diversity of swine IAVs. It has been mentioned before that pig movement within the USA, and between North America, Europe and Asia is associated with IAV gene flows (13, 24, 77). Viruses are inanimate microorganisms that cannot replicate outside host cells. Hence it is axiomatic that pig movement will be associated with regional IAV genetic diversity because infected animals that are transported are moving viruses with them.

Long-term persistence of IAV in populations has been reported in pigs before and after weaning and multiple reassortment events are well documented (17, 23, 37). In humans, people movement around the world is associated with IAV introductions to naïve populations and confirms that human movements and intermingling contribute to viral diversity (1). Nevertheless the roles of different pig subpopulations, and their movement within and among herds, on swine IAV diversity and evolution are not clearly understood. Furthermore, pork is the most consumed meat around the world. This high demand of pork has driven the industry, at least in the USA, to high efficiency farms with an average of 27 and 24 piglets born and weaned every year respectively (25-27).

Pig farms are mostly organized into multi-site production systems, where different phases of production are located at separate geographical sites. Typically, breeding, gestation and farrowing take place in breed-to-wean (breeding) herds. At approximately 21 days of age, piglets are weaned and transported to a nursery or a wean-to-finish site where pigs are raised until 10 weeks of age or to market-age, respectively. Breeding herds house replacement animals (gilts), adult females (sows), and piglets. Except on breeding farms where farrowing occurs in batches, typically piglets are born daily, with weekly births totaling about 40% of the resident adult female population. Adult females are replaced at a yearly rate of 45 to 55%, and replacement gilts are regularly introduced on schedules that range from 1 to 10 weeks or more across farms. As a result, breeding herd populations have high rates of turnover, which may be associated with fluctuating IAV susceptibility. For example, in a herd of 1000 sows that weans 25 pigs per sow per year at 3 weeks of age, the expected suckling piglet population is approximately 1440 (larger than the sow population), and around 450 piglets are born each week. The implications for IAVs evolution and emergence of new strains in this population undergoing rapid turnover are not fully understood, but the continual availability of new susceptible hosts with different levels of immunity to IAVs (acquired or maternally derived) may favor emergence of new IAV variants.

Table 1. Influenza A virus gene segments and proteins translated

Segment	Name	Number of Nucleotides*	Protein(s) translated	Function
1	Polymerase base 2 (PB2)	2316	PB2	RNA Polymerase complex
2	Polymerase base 1 (PB1)	2314	PB1 or PB1-F2	RNA Polymerase complex Endonuclease and elongation activity PB1-F2: pro-apoptotic
3	Polymerase acid (PA)	2250	PA	RNA Polymerase complex Protease Activity
4	Hemagglutinin (HA)	1776	HA	Surface antigen Binding and fusion
5	Nucleoprotein (NP)	1563	NP	RNA binding RNA synthesis Import
6	Neuraminidase (NA)	1410	NA	Surface antigen Neuraminidase
7	Matrix (M)	1030	M1/M2	Matrix protein Membrane Protein Nuclear exportation Budding, Ion channel Assembly activity
8	Non-structural (NS)	851	NS1/NEP	Multifunctional INF antagonist Nuclear exportation

*Based on the genome length of H1N1 A/IA/Sw/00239/04

Chapter 2: Association between influenza A virus infection and pigs subpopulations in endemically infected breeding herds

This work has been published in:

Diaz A, Perez A, Sreevatsan S, Davies P, Culhane M, Torremorell M (2015) Association between Influenza A Virus Infection and Pigs Subpopulations in Endemically Infected Breeding Herds. PLoS ONE 10(6): e0129213. doi:10.1371/journal.pone.0129213

Introduction

Influenza A viruses (IAVs) are Orthomyxoviruses able to infect many animal species including birds, pigs and humans (9). The segmented genome of IAVs allows the exchange of gene segments between IAVs during infection and replication (2) facilitating the emergence of novel IAV reassortants with pandemic potential. The 2009 pandemic IAV contained genes from swine IAVs circulating in North America and Eurasia (53) and highlighted the importance of pigs in the ecology of IAVs among species.

Influenza-like disease was first reported in pigs in 1918 at the time of the human Spanish flu pandemic and the virus was first isolated from pigs in 1930 (16, 107). Currently, IAV infections occur worldwide and are considered endemic in swine populations (107). In the US, IAVs have been present in pigs for many decades and several serological surveys conducted since the 1970's have consistently demonstrated that IAVs are ubiquitous in swine (108-110). IAVs remained genetically stable in pigs in the US with minimal or undetected viral evolution until 1998 when H3N2 reassortant viruses of swine, human and avian origin were detected in pigs (16). Subsequently new strains, new subtypes, and multiple reassortant viruses have been identified in pigs in North America (111, 112). There is evidence that multiple human introductions of IAVs, including the 2009 H1N1 pandemic virus, into the pool of IAVs circulating in the US have greatly contributed to the increase of genetic diversity of the virus (15, 112). The contemporary swine farming methods and live animal movements between farms and geographical regions increase IAV diversity (13, 24) and make IAVs harder to control in swine populations. However, limited information is available about the frequency of IAV introduction and its maintenance within swine herds.

IAV infections are common in the US Midwest, and herds can test positive year around with diverse viruses, regardless of the presence or absence of clinical

signs and vaccination status (17). Although influenza infections in individual pigs are of short duration (5-7 days), IAV infections of herds can be prolonged (weeks or months). When population dynamics, and extended IAVs circulation in pig farms are taken in consideration, the likelihood of reassortment increases (23, 37). Furthermore the introduction of IAV strains of human origin broadens virus diversity in pigs (61) and this diversity is accentuated by the frequent movement of weaned pigs into swine-dense areas (13).

The contemporary US swine industry is mostly organized into multi-site production systems. These systems have defined production stage facilities located in different geographical sites. Breeding, gestation and farrowing takes place in breed-to-wean (breeding) herds. At approximately 21 days of age, piglets are weaned and transported to a nursery or a wean-to-finish site where pigs are raised until 10 weeks of age or to market-age respectively. Breeding herds house replacement animals (gilts), adult females (sows), and piglets. Except on breeding farms where farrowing occurs in batches, typically piglets are born daily, with weekly births totaling about 40% of the resident adult female population. Adult females are replaced at a yearly rate of 45 to 55% (25), and replacement gilts are regularly introduced on schedules that range from 1 to 10 weeks across farms. As a result, breeding populations have high rates of turnover, and associated fluctuations in herd's susceptibility to IAV infections. Previous studies have shown that it is not easy to recover IAVs from sows (37, 113, 114). However, the dynamics of IAV infection among other subpopulations (gilts and suckling piglets) may influence the maintenance of IAVs in the breeding herd or contribute to the emergence of novel strains.

In this study, we aimed to define patterns of IAV infection in the most labile subpopulations on 5 IAV-infected breeding farms, with the goal of understanding the relative importance of these subpopulations as sources of new IAV infections in the breeding herd. This information is necessary to design targeted strategies

to control IAV infection within and between swine herds, and to minimize the risk of IAV infections of swine origin to people.

Materials and methods

Ethics statement:

Protocols and procedures followed throughout the study were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC 1207B17281), and the Institutional Biosafety Committee (IBC 1208H18341).

Study design and sample collection:

Five commercial pig-breeding herds (Farms 1 to 5) located in the Midwestern USA were conveniently selected for this study. Selection criteria included breed-to-wean herds with: a) confirmed IAV infection by real time reverse transcriptase polymerase chain reaction (RRT-PCR) within the past year, b) presence of an on-site gilt development unit (GDU), and c) introduction of external replacement animals into a defined isolation area. Each farm was visited monthly for 12 months and the overall sampling period spanned from November 2011 to December 2012. At each visit, three pig subpopulations were sampled: a) replacement females, resident on-farm for less than 4 weeks (new gilts), b) replacement females, resident on-farm for more than 4 weeks (gilts), and c) neonatal pigs less than 21 days of age (piglets). Adult animals (sows) were not included since prior studies consistently reported a low probability of influenza detection in this subpopulation (37, 113, 114). Due to schedules for delivery of replacement females, eligible populations of new gilts were only present at 21 of the 60 sampling events. A diagram of the study design can be seen in Fig 1. Information on management and control practices for IAV was recorded during the last visit.

Thirty pigs were selected from each pig subpopulation (new gilts, gilts and piglets) and sampled individually using a nasal swab (BBL CultureSwab, Becton Dickinson and Company, USA). Sample size (n=30) was estimated based on a 95% confidence to detect at least 1 positive sample if prevalence was 10% or higher at the subpopulation level. After sample collection, nasal swabs were refrigerated and transported to the laboratory on the manufacturer's transport media. At the laboratory, swabs were placed on 1.8 ml viral transport media (Dulbecco's Modified Eagle Medium (DMEM) plus 5% antibiotic-antimycotic, Gibco Life technologies, USA) vortexed for 10 seconds, alliquoted and stored at -80ºC until testing.

Influenza A virus detection and subtyping:

Nasal swabs were first screened for IAV by RRT-PCR on pools of three samples. If a pool tested positive then the samples comprising the pool were tested individually. All samples in a negative pool were considered negative. Viral RNA was eluted using 50 µl of each sample into 50 µl elution buffer using MagMax virus RNA isolation kit (Ambion, USA). Primers targeting the matrix (M) gene and AgPath-ID One-Step RT-PCR reagent kit (Ambion, Life technologies, USA) were used to detect IAV (84, 85). PCR mix containing 5 µl RNA, 12.5 µl 2X buffer, 1.0 µl 25X enzyme mix, 1.67 µl detection enhancer, 5 pmol of each primer and 1.5 pmol of probe was run on a LightCycler 480 system (Hoffmann-La Roche, Switzerland) at 45°C for 10 min, followed by 95°C for 10 min, and 45 cycles at 94°C for 1 sec and 60°C for 30 sec. Fluorescence was recorded at 60°C and a sample was considered positive if the cycle threshold (CT) was lower than 40. Positive samples with a CT value of 35 or lower were used for IAV virus isolation on Madin–Darby Canine Kidney (MDCK) cells (92) and each IAV isolate was subtyped based on the hemagglutinin and neuraminidase.

Data analysis:

Data collected in the survey was summarized by farm and three independent variables were taken into consideration for statistical analysis: a) pig subpopulation, b) farm, and c) annual quarter (1^{st} to 4^{th}). The association between IAV detection and subpopulation was analyzed at the group level where a group was defined as the subpopulation of pigs that was sampled during a given visit and considered positive if one or more swabs within the group tested positive to IAV. A Pearson's Chi square or Fisher's exact test was used to compare the frequency of IAV positive and negative outcomes by subpopulations (new gilts, gilts and piglets), farms (1 to 5), and annual quarter. Using gilts as the reference group, an unconditional logistic regression model was used to measure the crude association between the outcome and pig subpopulations, farm and annual quarter. Finally, fixed and mixed logistic regression models were compared to estimate the associations between IAV infection and subpopulation adjusting by farm, annual quarter, and sampling event. Subpopulation and annual quarter were included as fixed effects in the models, and clustering variables (e.g. farm and farm visit) were included as random effects. Given the study design (Fig 1) we assumed that samples clustered by sampling visit, farm, and annual quarter. We also assumed that pig subpopulations were nested within sampling visits, and that sampling visits were nested within farm. A Wald Chi-square was used to test the significance of individual coefficients within each model and a likelihood ratio Chi square test was used to compare hierarchical models. Statistical significance was assumed at p values lower than 0.05 and the model with the lowest Akaike information criterion (AIC) value was selected as the final model. All data analysis and graphics were performed using R 3.1.0 (The R Foundation for statistical Computing, www.R-project.rog) and packages installed included lattice (115), gmodels, car, aod (116), ggplot (117) , lme4 (118) and psych (119).

Results

Between November 2011 and December 2012, 4190 individual nasal swabs were collected from 141 groups of pigs in the 5 pig-breeding herds. Farm demographics are summarized in Table 2. Among all groups sampled, 60 (42.5%), 60 (42.5%) and 21 (15%) were groups of piglets, gilts and new gilts respectively. Swabs were collected from 1796 (43%) piglets, 1768 (42%) gilts and 626 (15%) new gilts. Although all farms and subpopulations under study tested positive to IAV at least once, only 28 groups (19.9%) and 324 swabs (7.7%) were positive for IAV. Piglets tested positive at least once in all farms, new gilts only tested positive in farms 1 and 3, and all gilts tested negative in farm 5 (Fig 2). One hundred and twenty four IAVs isolates were recovered and 123 of them were successfully subtyped. Subtypes H1N1, H1N2 and H3N2 were identified and more than one IAV subtype was isolated in all farms over time.

No positive samples were found in May, June or September 2012 and the univariate analyses indicated that IAV infection was strongly associated ($p<0.01$) with pig subpopulation and annual quarter (Table 3). The crude odds of IAV infection in groups were higher for piglets and new gilts compared to groups of gilts, and lower for groups of pigs sampled during the second annual quarter (April, May, June) compared to any other quarter of the year (Table 4).

The final multivariate model indicated that, after adjusting by annual quarter and sampling visit, the odds of IAV infection were higher in groups of new gilts (OR=7.9 95% CI: 1.4,43.9) and piglets (OR=4.4 95% CI: 1.1,17.1) compared to groups of gilts (Table 5). Inclusion of farm to the model was not statistically significant ($p>0.05$) therefore this variable was excluded from the final model.

Discussion

To advance our understanding of IAV epidemiology in swine breeding herds, we combined frequent sampling and PCR-based methods to define patterns of active IAV infections among pig subpopulations present in these herds. We found that replacement animals resident on-farm for less than 4 weeks (new gilts) and pigs less than 21 days of age (piglets) had higher odds of testing positive to IAVs compared to replacement animals resident on-farm for more than 4 weeks (gilts). Therefore new gilts and piglets may represent the most epidemiologically significant reservoirs for IAVs in swine breeding herds. Sows were not included in our study because they have been found to have a low probability of influenza positivity in endemically affected herds (37, 113, 114). Our results also indicate that there was a strong association between IAV infection and annual quarter and that this association was still statistically significant after controlling for the subpopulation effect.

In the USA there are approximately 65 million commercial pigs, of which approximately 6 million are breeding sows. Breeding females are replaced at a 45-55% rate annually (i.e. approximately 3 million gilts are introduced every year into US sow farms to replace existing breeding stock), and each sow gives birth and weans approximately 27 and 24 piglets respectively (25). For example, in a herd of 1000 sows that weans 25 pigs per sow per year at 3 weeks of age, the expected suckling piglet population is approximately 1440 (larger than the sow population), and around 450 piglets are born each week. The implications for IAVs evolution and emergence of new strains in this population undergoing rapid turnover are not fully understood, but the continual availability of new susceptible hosts with different levels of immunity to IAVs (acquired or maternally derived) may favor emergence of new variants.

Most published studies to date have studied IAV transmission at a broad scale based on the genetic evolution of IAV throughout time and space (13, 15, 24,

61), and relatively limited information is available on the epidemiology of IAVs at the herd level. Furthermore, most herd level studies have used serological methods rather than direct detection of the virus by molecular methods (120-122). Serological results are less definitive and can be difficult to interpret since detected antibodies may reflect maternally derived antibodies or active immunity to IAV infection or vaccination, none of which can be distinguished from the other.

This study demonstrates that in certain breeding herds, new gilts and piglets can be an important reservoir for IAVs in swine populations. New gilts represented animals from an external source and could have been naïve to resident farm viruses or a source of new IAVs to the breeding herd. Upon arrival to a farm, new gilts were commonly kept in a separate room or building (although rarely in complete isolation) to minimize introduction of new diseases for approximately 30 days. After that, gilts were moved into the gilt development unit. We selected 4 weeks as the cut-off to classify gilts (new gilts or gilts) to reflect this industry practice of gilt management. Our results then showed that gilts had lower odds of IAV infection. One possible explanation is that gilts have been able to clear the infection since they have been on site for a longer period of time. IAV infections at the individual animal level are self-limiting and usually last between 5 to 7 days (16). In contrast, finding IAVs more frequently in new gilts may reflect the introduction of infected animals from the source herds. Although in this study we did not sample the source herds, the likelihood of these herds to be IAV positive is high given the commonality of IAV infections in the Midwest (17). However, we cannot fully rule out the possibility that new gilts became infected with resident viruses after arrival to the herd. Although new gilts were placed into isolated designated areas and procedures were in place to minimize disease transmission (eg. isolation, vaccination), these areas or procedures might not have been able to fully contain infections within the designated areas. Future studies sampling gilts at arrival to the breeding herd would be required to more precisely estimate the risk of gilts at introducing IAVs into breeding herds.

Nevertheless our study identified new gilts as a subpopulation that tested IAV positive more frequently than resident gilts which indicates the need to have specific control programs or protocols to mitigate their risk.

In contrast, neonatal piglets at birth are immunologically naïve and likely acquire IAV infection within the breeding herd. The role of piglets in maintenance and dissemination of IAV infections has been documented before (37). In this study we support those findings although our study design did not discern whether infections in piglets in consecutive months were due to endemic viruses maintained in this subpopulation, or whether they represented new infections from other subpopulations. Previous studies have shown that different IAV subtypes can co-circulate, be sustained (17) and reassort (23) in swine populations. In this study we also demonstrate that different IAV subtypes can be present simultaneously in breeding herds. Further studies on the complete genome characterization and phylogenetic analysis of the viruses recovered from these subpopulations will be able to assess virus diversity and reassortment within these pig subpopulations. Additionally, our study measured the odds of IAV infection based on the detection of the matrix gene. The matrix gene segment is highly conserved among IAVs. Whether the odds of IAV infection in pig subpopulations varies for different IAV subtypes should also be further investigated.

Time of the year when IAV infection was detected was also associated with IAV infection in our study. In humans, IAV infections have a clear seasonal pattern in temperate regions but the pattern is less defined in the tropics (43). In swine, IAVs seasonality is still under debate (17, 75, 78). One recent study in Europe did not detect any seasonal trend of swine IAV (18). In another study, season (winter) and IAV like-illness in pigs were strongly associated but IAVs detection in pigs and season were not (78). In our study, we found higher odds of IAV detection in groups of pigs sampled during the first quarter of the year (winter in the northern hemisphere), thus favoring the hypothesis of a seasonal trend of

swine IAV. The differences among studies may be due to study design, sample size, and the result of measuring these associations using active or passive surveillance.

We acknowledge that our results do not represent IAV infection dynamics across U.S herds since only a limited number of herds participated in the study and these herds were conveniently selected. In addition, we identified significant variability in the number of IAV infected pigs between sampling events, which may be a reflection of the sample size or number of farms in the study. However, our IAV detection rates are similar to those described in other studies and our sample size for each of the farms is larger than in other studies (3).

In conclusion, our study indicates that there are differences in the odds of IAV infection across different pig subpopulations found in breeding herds, and that IAVs can be found more frequently in new gilts and piglets than in resident gilts. Our results also indicate that there is a strong seasonal component to IAV infection in breeding herds from this study with the first quarter of the year being the period with most positives. Overall our results contribute to better understand IAV transmission in pigs and indicate the need to focus interventions to control IAV infections in new gilts and piglets before weaning.

Table 2. Descriptive Farm Demographics.

Farm	1	2	3	4	5
State	South Dakota	Minnesota	Iowa	Minnesota	Minnesota
Date when sampling started	Jan-12	Jan-12	Nov-11	Nov-11	Jan-12
Sow inventory	3420	2343	3249	1000	3000
Sow vaccination against IAV	No	Yes	Yes	No	No
Average age of piglets at weaning (days)	19	20	20	18	21
Frequency of gilt introduction (weeks)	5	5	5	20	5
Average age of new gilts at arrival at the farm (weeks)	21	3	3	16[1]	22
Vaccination of gilts after arrival	No	Yes	No	Yes	Yes
Respiratory clinical disease in piglets	Yes	Yes	No	Yes	No
Respiratory clinical disease in new gilts	No	No	Yes	No	Yes

[1]Age range for new gilts at arrival in farm four was between 8 and 24 weeks. For all other farms it was one week.

Table 3 Number (%) of pig groups positive to influenza A virus by RRT-PCR distributed by subpopulation, farm, and annual quarter.

A group was considered positive if at least one swab within the group was IAV RRT-PCR positive. The proportion of positive vs. negative outcomes by subpopulation, farm, and annual quarter was compared using a Pearson chi-square test (or Fisher's exact if needed) and the p-value for each comparison is shown.

Variable		Positives (%)	Negatives (%)	Total
Subpopulation (p<0.01)	New gilts	8 (38.1)	13 (61.9)	21
	Gilts	6 (10.0)	54 (90.0)	60
	Piglets	14 (23.3)	46 (76.7)	60
	Total	28 (19.9)	113 (80.1)	141
Farm (p=0.17)	1	9 (32.1)	19 (67.9)	28
	2	3 (10.7)	25 (89.3)	28
	3	9 (27.3)	24 (72.7)	33
	4	4 (15.4)	22 (84.6)	26
	5	3 (11.5)	23 (88.5)	26
	Total	28 (19.9)	113 (80.1)	141
Annual quarter (p<0.001)	1st	16 (40.0)	24 (60.0)	40
	2nd	1 (3.0)	32 (97.0)	33
	3rd	3 (8.8)	31 (91.2)	34
	4th	8 (23.5)	26 (76.5)	34
	Total	28 (19.9)	113 (80.1)	141

Table 4. Results from the univariate analysis.

The crude association between IAV detection and subpopulation or annual quarter was measured through odds ratios. A group was considered positive if one or more swabs within the group tested positive to IAV by RRT-PCR. The first group for each variable of interest was used as the reference group.

Variable	Group	OR (95% CI)
Subpopulation	Gilts	-
	New gilts	5.5 (1.6, 19.6)***
	Piglets	2.7 (1.01, 8.3)**
Annual quarter	2nd. Apr, May, Jun	-
	3rd. Jul, Aug, Sep	3.1 (0.4, 64.5)
	4th. Oct, Nov, Dec	9.8 (1.7, 188.8)*
	1st. Jan, Feb, March	21.3 (3.9, 398.4)**

*p<0.05, **p<0.01, ***p<0.001

Table 5. Results from the multivariate analysis (Mixed effects model).

While subpopulation and annual quarter were included as fix effects, sampling visit was included as random effect.

Variable	Group	OR (95% CI)
Subpopulation	Gilts	-
	New gilts	7.9 (1.4, 43.9)*
	Piglets	4.4 (1.1, 17.1)*
Annual quarter	2nd. Apr, May, Jun	-
	3rd. Jul, Aug, Sep	3.5 (0.2, 54.9)
	4th. Oct, Nov, Dec	16.1 (1.1, 234.7)*
	1st. Jan, Feb, March	43.9 (2.8, 686.8)**

*p<0.05, **p<0.01, ***p<0.001

Figure 1. Study design.

Five swine-breeding herds (largest rectangles, 1 to 5) in the Midwest were sampled between November 2011 (month 1) and December 2012 (month 14). White rectangles represent sampling visits (12 per farm), and the smallest rectangles (n=141) indicate the groups of pigs sampled. Groups are colored based on pig subpopulation: yellow (new gilts, n=21), green (gilts, n=60), and purple (piglets, n=60). Missing group-rectangles indicate that there were no new gilts on that visit. Groups were assumed nested within sampling events, and sampling events were assumed nested within farms.

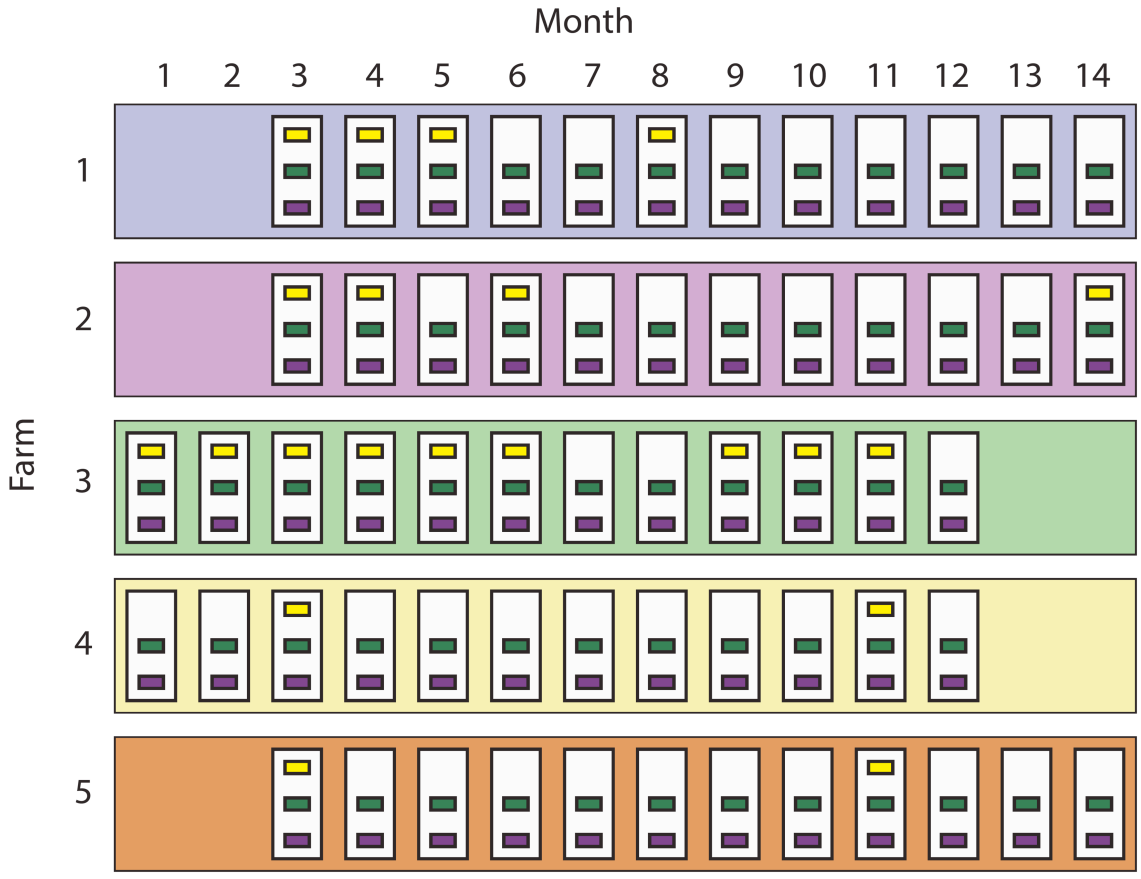
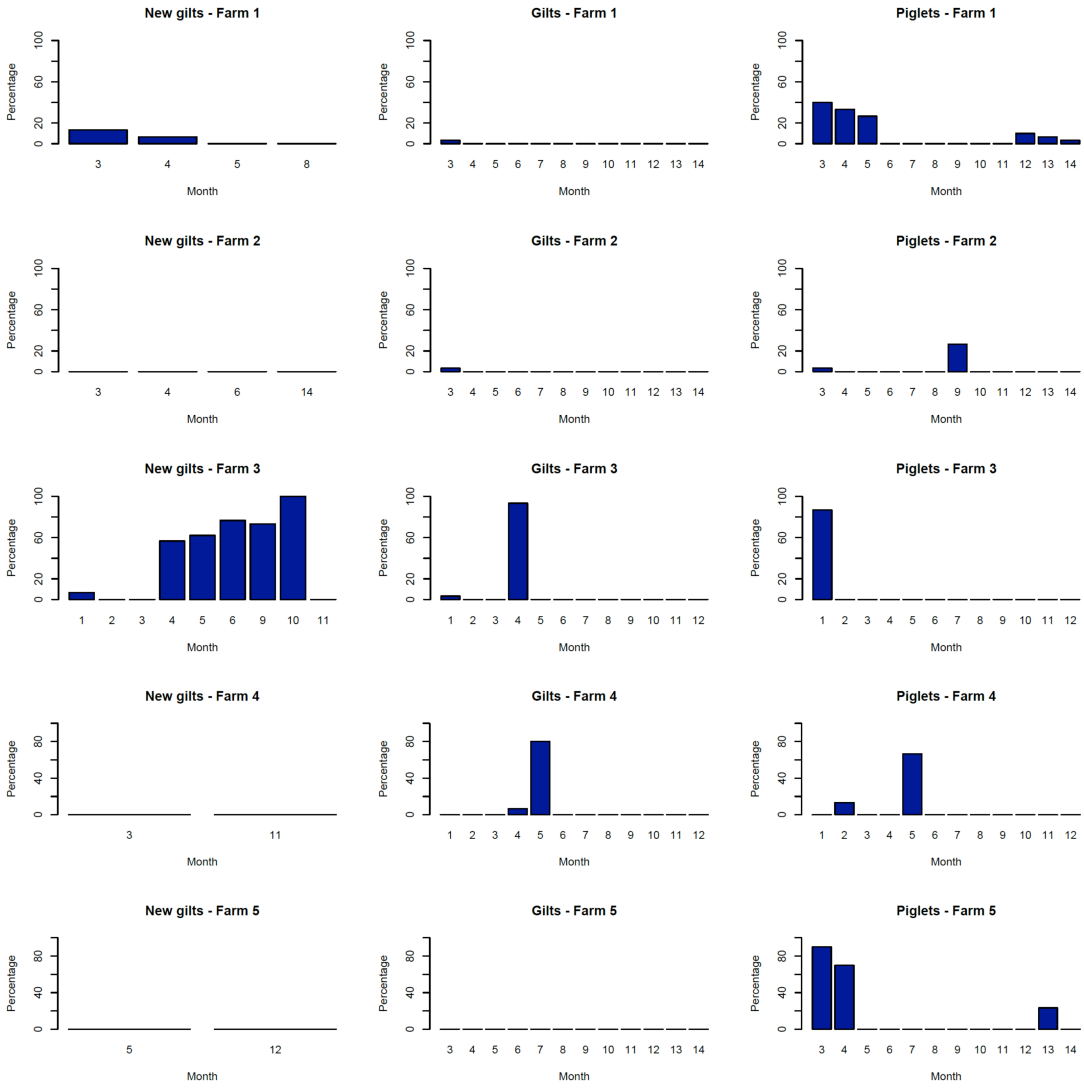


Figure 2. Percentage of influenza A virus positive samples distributed by farm, subpopulation and month.



Chapter 3: Deep genome sequencing of Influenza A viruses in pig breeding herds reveals the emergence, persistence, and subsidence of diverse viral genotypes over time.

Introduction

In 2009, a novel influenza A virus (IAV) that emerged from swine caused the first human IAV pandemic of the 21st Century and changed perceptions of the role of pigs in the ecology of IAV infections (10). IAV infections are endemic in many species including wild waterfowl (123), humans (43), and pigs (16, 86). IAVs can also infect poultry (124, 125), horses (126), cats (6), dogs (5), some marine mammals (7) and a distant genetic lineage of IAV has been recently identified in bats (8, 45). Zoonotic IAVs can cause pandemic infections (10), however not all zoonotic IAV infections exhibit sustained human-to-human transmission (12, 47). Except for the 2009 pandemic, no other swine-origin transmission of IAV has acquired the ability to transmit effectively human-to-human; however, several reverse zoonotic events have resulted in human-IAVs to becoming well established in swine populations (13, 15, 61).

IAVs are Orthomyxoviruses with eight single stranded negative sense RNA gene segments that translate at least 12 different proteins (9). IAV genome segments include polymerase B2 (PB2, segment 1), polymerase B1 (PB1, segment 2), polymerase A (PA, segment 3), hemagglutinin (HA, segment 4), nucleoprotein (NP, segment 5), neuraminidase (NA, segment 6), matrix (M, segment 7), and non-structural protein (NS, segment 8). RNA viruses change rapidly over time (44), and IAVs are no exception. However, not all IAV gene segments change at the same rate and substitution rates can be host-species specific (59).

Additionally, the segmented nature of the IAV genome allows two or more IAVs to exchange gene segments (reassort) during replication (70, 71) increasing factorially the potential for virus diversification. IAVs are commonly found in North American swine herds (37, 113, 114) where H1N1, H1N2, and H3N2 subtypes are frequently identified (17). At the HA level, these IAVs cluster in six antigenically and phylogenetically distinct H1 clades (α, β, γ1, γ2, δ1, and δ2) (55, 76) and four H3 clusters (I, II, III and IV) (54). The genetic diversity of IAVs in United States (US) pigs has been associated with the geographical distribution and movement of pigs (13) and the international trade of pigs is considered a predictor pathway of genetically-related IAVs being introduced from the US and Europe to Asia (24, 77).

In the contemporary US swine industry, there are almost 70 million pigs, which include 6 million breeding stock. Breeding, gestation and farrowing takes place in breeding herds (127). Then, at about 21 days of age piglets are commonly weaned and transported to a separate site to be reared to the next production phase or until market (25). Pig breeding herds house different types of pig subpopulations with different turnover rates and different susceptibilities to IAV infection (128). These sub-populations include sows (mothers of piglets), replacement animals for the sows (gilts), and piglets (pigs from birth to weaning). Gilts are introduced on a regular basis to the breeding herd and replace sows at a yearly rate of 45 to 55% (25). Moreover, suckling piglets represent approximately 40% of the resident adult female population and every week

newborn piglets replace those pigs that are being weaned. Therefore breeding herds are dynamic populations and suckling piglets represent the largest population and have the highest turnover rate among all different pig subpopulations present.

Despite all the knowledge gained on IAV diversity in pigs as a result of the increased surveillance efforts of the last few years, there is limited information on the genetic diversity and virus evolution observed at the herd level. We hypothesize that the population dynamics present in pig breeding herds play a key role in the introduction, emergence, and persistence of IAV over time. Different pig subpopulations, with their respective turnover rates, may represent different ecological niches for IAV replication and evolution. Understanding the evolution and diversity of IAVs among swine subpopulations in breeding herds is crucial to unravel the mechanisms by which IAVs persist for prolonged periods of time in these herds; therefore, we characterized the complete genome of IAVs during infection of pigs under natural conditions and demonstrated the complex and dynamic occurrence and maintenance of IAVs in pig subpopulations that are present in pig breeding herds. Results from this study provided a deeper understanding of IAV persistence at the herd level, knowledge that is required to design more effective health interventions to control IAV infection in pigs and reduce its zoonotic potential.

Materials and methods

Ethics statement

Protocols and procedures followed throughout the study were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC 1207B17281), and the Institutional Biosafety Committee (IBC 1208H18341).

Study design, IAV detection and isolation:

We designed a longitudinal study with multiple cross-sectional sampling events to detect and characterize IAV infections in five commercial pig breeding herds (Farm 1 to 5) located in the Midwestern, USA. More specifically we evaluated at the farm level the emergence (first isolation of an IAV), persistence (the same IAV lineage detected over time), and subsidence (no further isolation of an IAV previously recovered) of IAVs in these five herds. All farms had a history of IAV infection and were conveniently selected on November 2011 and sampled on a monthly basis for 12 months. During each farm visit, 30 nasal swabs (BBL CultureSwab, Becton Dickinson and Company, USA) were collected from three pig subpopulations: 1) new gilts (replacement breeding stock on farm for less than 4 weeks), 2) gilts (replacement breeding stock on farm for more than 4 weeks) and, 3) piglets (3 week-old suckling pigs). Sows were not sampled because previous studies have found that recovering IAVs from sows is frequently unsuccessful (37, 113, 114). Sample size (n=30) was calculated at the subpopulation level with

a 95% confidence level to detect at least 1 positive sample if the prevalence was 10% or higher. Overall, new gilts were sampled during 21 visits due to varying schedules of the delivery of replacement animals. Once collected, swabs were refrigerated and transported to the laboratory on the manufacture's transport media and then placed into 1.8 ml sample storing media (Dulbecco's Modified Eagle Medium (DMEM), 2% bovine serum albumin (BSA) fraction V 7.5% solution (Gibco, Life technologies, USA), 5% antibiotic-antimycotic (Gibco, Life Technologies, USA containing 10000 IU/ml of penicillin, 10000 µg/ml of streptomycin, and 25 µg /ml of Fungizone)). Swabs in the sample storing media were vortexed for 10 seconds, and then stored at -80ºC until IAV testing.

Samples were initially tested for IAV in pools of three by reverse transcriptase real time polymerase chain reaction (RRT-PCR) targeting the matrix gene using methods described elsewhere (84, 85). Each pool contained only samples from the same farm, month and subpopulation. If a pool tested positive then aliquots of the original samples were tested individually. A test was considered positive when the RRT-PCR cycle threshold (Ct) value was lower than 40 and IAV isolation was attempted from all swabs with a Ct value $\leq$ 35. Madin–Darby canine kidney (MDCK) epithelial cells were used for IAV isolation (92). Briefly, one six well plate (Corning, Sigma-Aldrich, USA) was used per sample to avoid cross contamination between samples and two negative controls were used per plate. When the cell monolayer was ~90% confluent the cell growth media was discharged and then each well was washed twice with Hank's solution (Gibco,

Life technologies, USA) containing 0.15% 1mg/ml TPCK trypsin (Sigma-Aldrich, USA). 200 µl of 1:1 and 1:2 dilution of the sample were used in replicates to infect 4 wells of each plate and the two negative controls were mock infected with 200 µl of DMEM (Gibco, Life technologies, USA). Plates were placed into a 5% CO_2 incubator for an hour and then 2 ml of viral growth media was added to each well. Viral growth media contained DMEM (Gibco, Life technologies, USA), 4% bovine serum albumin (BSA) fraction V 7.5% solution (Gibco, Life technologies, USA), 0.15% 1mg/ml TPCK trypsin (Sigma-Aldrich, USA), and 1 % antibiotic-antimycotic (Gibco, Life Technologies, USA). Plates were observed daily and harvested if IAV cytopathic effect (CPE) was visually confirmed. If no CPE was present, then wells were harvested at 7 days post infection for a blind passage in a new MDCK plate. A hemagglutination assay was performed on all wells harvested. If a well lacked CPE but was hemagglutination positive, a blind passage was performed on a new MDCK plate. IAV isolation was confirmed by CPE and antigen detection using swine influenza type A antigen test kit FluDetect (Zoetis, USA). Initial IAV positive isolates (passage 1) were expanded into T25 flasks (passage 2) and these second passages were used for complete genome amplification and sequencing.

Complete genome amplification and sequencing:

The complete genome of IAV was amplified in a single reaction as previously described (103). IAV RNA was extracted from positive isolates using MagMax

Viral RNA isolation kit (Ambion, Life Technologies, USA). RRT-PCR was achieved using SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen, Life Technologies, USA). A 50 µl PCR mix was prepared containing 10 µl DNase/RNase-Free distilled water (Gibco, USA), 25 µl 2x reaction mix, 1 µl SuperScript III RT mix, 1 µl (10 µM) of each primer (MBtuni12(M): ACGCGTGATCAGCRAAAGCAGG and MBtuni13: ACGCGTGATCAGTAGAAACAAGG), and 12 µl of RNA template. PCR products were verified by gel electrophoresis, purified using QIAquick Spin Kit (QIAGEN, USA), eluted in 20 DNase/RNase-free distilled water (Gibco, Life Technologies, USA) and submitted for next generation sequencing (NGS) using Illumina MiSeq system (Illumina, USA) at the University of Minnesota Genomics Center (UMGC).

Sequencing data was analyzed through the resources available at the University of Minnesota Supercomputing Institute (MSI). Sequencing quality was first verified using FastQC (129) and then trimmed using the pair-end mode of Trimmomatic (102). Sequencing assembly was performed using Bowtie2 (130) and SAMTools (97) on a reference template containing 6 IAV internal gene segments (PB2 (CY099076.1), PB1 (CY099309.1), PA (CY045233.1), NP (CY009919.1), M (DQ150436.1), and NS (CY050162.1)) and 4 antigenic gene segments (H1 (FJ789832.1), H3 (KC992248.1), N1 (GU236519.1), N2 (KC866483.1)). Consensus sequences for each contig assembled were trimmed to coding IAV gene regions and their functionality verified using the NCBI FLu Annotation web-service (FLAN (131)).

Phylogenetic origins and IAV diversity within and between farms:

First, all IAV isolates from this study were classified based on the HA and NA and then specific phylogenetic relationships were estimated for each gene segment. Data sets for this analysis included the sequences from this study plus all complete swine IAV sequences from the USA (gene segments 1 to 8) downloaded from the Influenza Research Data Base (IRD (132)) on October 16th 2014 and an additional data set from the United States Department of Agriculture (USDA), National Veterinary Service Laboratories (NVSL). Only IAVs collected between January 1st 2003 and October 16th 2014 with collection date (month/day/year) were kept, and 01/01/03 was set as day 1 for purposes of evaluating IAV evolution. Data sets for antigenic HA gene segment subtype H1 was broken accordingly to Zell et al., 2013 (133) into lineage 1A (H1 classical-swine), 1B (human seasonal H1). All other antigenic data sets (H3, N1, and N2) were kept as single genetic lineages. Additionally, data sets for segment 7 (matrix) and 8 (non-structural) were broken into pandemic and non-pandemic sequences for a total of 13 IAV gene segment datasets. Each dataset was aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE (134)) and approximately maximum-likelihood trees were constructed for each dataset using FastTree2 (135), assuming a Generalized Time Reversible (GTR) substitution model. Local support values were estimated under the discrete gamma model (136) with 20 rate categories (Gamma20-based likelihood). The

best fitting root for each tree was found using Path-O-Gen v1.4 and a simple linear regression model was used to estimate the crude association between distance to the root of the phylogenetic tree and time (in days). Furthermore, the line fitted was used to estimate the x-intercept (distance to the root equal 0) and obtain a crude estimate of the time to most recent common ancestor (TMRCA) for each genetic lineage analyzed.

The temporal distribution of IAV was compared within and between farms based on the HA lineage, and reassortment events were evaluated using the HA as a backbone. Additionally, sequences obtained from this study were aligned using ClustalX (137) and the pairwise distance between sequences used to estimate the genetic diversity of IAV isolates over time. Finally, hypothetical HAs and NAs were translated and polymorphic amino acids inferred. Briefly, HA and NA amino acid sequences were first aligned using ClustalX (cost matrix Blosum62) and then polymorphic sites were stripped to estimate the frequency of each amino acid. All graphics were performed using R 3.1.0 (The R Foundation for statistical Computing) and packages installed included gplots, RcolorBrewer, and heatmaps.2.

Results

One hundred and twenty four IAVs were isolated from 207 IAV RRT-PCR positive nasal swabs (Table 6) and the complete IAV genome was successfully

sequenced from 123 of these virus isolates. IAV isolates were obtained from new gilts (farm 3), gilts (farm 3 and 4) and piglets (farms 1-5). While H1N1 and H3N2 viruses were isolated from farms 1-5, H1N2 viruses were only isolated from farms 2, 3 and 4. Ninety seven percent of virus isolates (n=120) contained a single IAV subtype with either H1N1 (n=31, 25.2%), H1N2 (n=26, 21.1%), or H3N2 (n=63, 51.2%) subtypes. However, three (2.5%) IAV isolates had a mixture of IAV subtypes and those were recovered from two gilts in farm 3 (H1N1/H3N2 and H1N1/N2 respectively) and one piglet in farm 2 (H1N2/H3N2). Additionally, co-circulation of two or more IAV subtypes at one sampling period was confirmed at least once in all farms

To understand the phylogenetic relationships of these 123 viruses, their gene segment sequences (n=1000) were compared to 14,401 sequences from IAVs circulating in the US between January 1st 2003 and October 16th 2014. For all swine IAV gene segments analyzed (Table 7) a linear association was present ($p<0.05$) between the distance-to-the-root of the phylogenetic tree and time (in days). Furthermore, 51% to 86% of the variability on the distance-to-the-root was explained by time. The nucleotide difference between sequences ranged between $4.64x10^{-6}$ (NS gene, segment 8) and $1.51x10^{-5}$ (NA, subtype N2, segment 6) for each day increment over time. Moreover, the 125 HA sequences from this study clustered within 7 different monophyletic clades (Gamma20 based likelihood > 0.91 (min: 0.91 max: 0.99)). For this study, these seven HA clades were denominated study clades 1-7 (SC1 to SC7) and are illustrated in

Figs 3a, 3b and 3c. Study clades 1, 2 and 3 were H1 viruses although SC1 viruses clustered with swine IAV gamma viruses while SC2 and SC3 clustered with swine IAV delta viruses (55, 56, 76). Furthermore, IAVs study clade 4, 5, 6, and 7 (Fig 3c) were all H3 viruses and clustered with H3 cluster IV (54).

The remaining gene segments were classified based on these 7 HA study clades. The distribution of the neuraminidase gene segment among IAVs circulating in US swine is illustrated in Fig 3d (subtype N1) and Fig 3e (subtype N2). Most of the internal gene segments (1, 2, 5, 6, 7, and 8) clustered also into seven monophyletic clades (Fig 4); however eight sequences from these gene segments (highlighted with an asterisk in Figs 3 and 4) did not cluster within the expected phylogenetic clade. Additionally, there were eight monophyletic clades for PA gene (segment 3) because IAVs in HA SC7 had two different PA lineages (Fig 4c, named 7a and 7b). These genome differences allowed us to identify 12 reassorted IAVs (Fig 5) during the study period. Reassorted IAVs were found in farms 2, 3, and 5 and in the three pig subpopulations studied.

Viruses with HAs clustering in study clades 1, 2, 3, 4, and 7 had matrix genes (segment 7) of pandemic origin (Fig 4e), while viruses with HAs clustering within study clades 5 and 6 had matrix genes of non-pandemic origin (Fig 4f). Furthermore, 79% (n=98) of the matrix gene sequences (segment 7) from our study contained a signature mutation (S31N) that has been associated with resistance to the antiviral amantadine (138). In contrast, viruses with HAs

clustering in clades SC3 and SC7 had non-structural genes (segment 8) of pandemic origin (Fig 4g) while the remaining viruses carried non-structural genes of non-pandemic origin (Fig 4h).

To better understand the within farm diversity of IAVs in pig breeding herds, the pairwise percent nucleotide identity (ClustalX alignments) was estimated for each segment sequenced during the study period. Detailed results for HA and NA comparisons are shown in Fig 6. Nucleotide sequence identity for PB2, PB1, PA, HA subtype H3, NP, NA subtypes N1 and N2, and NS segments was higher than 90%. However, the minimum pairwise percent identity for segments HA subtype H1 and M were 74.7% and 86.9%, respectively. Furthermore, the temporal distribution of IAVs by clade, farm, and pig-subpopulation is shown in Fig 7 and illustrates the emergence, persistence and subsidence of different HA lineages over time. At the subpopulation, level only piglets and new gilts harbored IAVs from more than one HA study clade at a given sampling visit. IAVs clustering within three different HA study clades were isolated from piglets and new gilts in farm 3 at month 1 and 10 respectively, and from piglets in farm 2 at month 9. Moreover, IAVs from study clades 2, 3, and 6 were recovered only once over time within the same farm while IAVs from study clades 1, 4, 5, and 7 were isolated multiple times (Fig 7).

While IAVs from SC1 persisted in multiple farms (farms 1, 3 and 4) IAVs from study clades 4, 5, and 7 persisted only in farms 1, 5, and 3, respectively (Fig 7).

At the farm level, the minimum HA percentage identity among persistent IAVs was 98% and the time of persistence ranged between 35 and 283 days (Fig 8). Moreover, IAVs from study clades 1 and 7 persisted simultaneously in farm 3 (Fig 7). Initially, IAVs from study clades 1, 3 and 7 were co-circulating in piglets at month 1 in farm 3; then, IAV from SC3 were not isolated again, and only study clade 1 and 7 viruses were recovered at a later point from gilts at month 4, and from new gilts at months 5, 9, and 10.

To expand our understanding on the molecular diversity of IAVs in pig breeding herds we translated the HA and NA genetic sequences into hypothetical protein sequences and compared them. We stripped the polymorphic amino acid sites from the HA and NA hypothetical-proteins alignment and compared them among HA study clades (Fig 9). The number of polymorphic amino acid sites among HA study clades in the HA1 region ranged between 1 and 15 (Fig 9a). No polymorphic amino acids were found within the HA2 region of HA and only three polymorphic sites were found within the signal peptide region (C12Y, in sequences from clade 5; Q17L, in sequences from clade 6; and S8N in sequences from study clade 7). Only one site in HA proteins from study clade 5 (amino acid 503) had 3 polymorphic amino acids at these positions (E, D and N) while all other sites had only 2 polymorphic amino acids. Finally, the number of polymorphic sites among NAs ranged between 1 and 29 (Fig 9b) and only one site in NA proteins from viruses in SC7 (amino acid 370) had 3 polymorphic amino acids at this position (I, L and S).

Discussion

To evaluate the epidemiology and genetic diversity of IAVs in pig breeding herds over time, we conducted a one-year longitudinal study and characterized the complete genome of 123 virus isolates recovered from three pig subpopulations in 5 Midwestern pig breeding herds from the USA. Our results showed evidence of emergence, maintenance and subsidence of different viral genotypes over time. Furthermore, we revealed the dynamic nature of co-circulating IAV genotypes within farms and the on-going exposure of pigs to different antigenic subtypes and antigenic viral variants within subtypes. Our findings illustrated the complexity of IAV genetic diversity within breeding herds and among pig sub-populations and provide a better background for IAV control in pig breeding herds.

Understanding IAV evolution in endemically infected pig-populations is complex due to the molecular characteristics of the virus and the high turnover rates of the host. In our study, multiple IAVs co-existed at the herd level in one or multiple pig subpopulations. However, some IAV genetic lineages persisted over time while others appeared to “subside” or disappear. The persistence of IAVs in pig farms has been reported before (17) and facilitates virus drift over time. Nevertheless, the extent of genetic analysis performed in this study provided information on genetic and antigenic diversity of IAVs in endemically infected breeding herds to

a level that has not been previously available as well as a more accurate characterization of the dynamics of IAV diversity within swine herds. Our IAV isolates included genotypes that were closely related to each other (e.g. viruses clustering within the same phylogenetic clade) and viruses that were clearly distinct (e.g. different IAV subtypes) and illustrated the dynamic makeup of IAV genome in pig breeding herds.

The ecological circumstances that allow IAVs to persist or disappear in swine populations are still not clearly determined. The simplest explanation for this dynamic occurrence of IAVs in swine populations is the antigenic diversity of the virus. IAVs can drift over time avoiding immune responses against divergent IAVs (41, 62, 82). At the HA level, swine IAVs with different genetic lineages are known to have different antigenic properties (76, 139). Furthermore, in pigs with or without passive immunity to IAVs, HA nucleotide mutations can happen rapidly after infection (140). However, IAVs are transmitted with a collection of HA1 alleles (sequence variants) that can emerge or disappear during infection of naïve or immune pigs (73). This dynamic transmission of different IAV alleles during infection of vaccinated pigs involves the entire genome of the virus and not only HA (141), which suggest that the emergence, maintenance and subsidence of certain IAV genotypes over time at the population level might start with varying virus replication rates at the individual level specifically due to varying host immune responses (82, 142, 143).

In this study we focused on those pig subpopulations with the highest turnover rates in pig breeding herds because these subpopulations are expected to have different ages and diverse immune statuses. In mice, different immune status (naïve vs. vaccinated) has been associated with antigenic drift of IAVs (63) although one study comparing naïve and vaccinated pigs did not find differences in the genetic diversity of IAVs between groups during infection (73). While the antibody repertoire of piglets (3 weeks or younger) is mostly maternally derived; in gilts and new gilts this repertoire depends basically on previous exposures to IAV antigens. Hence, future studies addressing the diversity of maternally derived immunity in piglets before weaning and its effect on IAV diversity and evolution are required.

Moreover, in the USA there are almost 70 million commercial pigs, which include around 6 million sows. Every year, gilts replace 45 to 55% of the sow population and each sow gives birth and weans 27 and 24 piglets respectively (25). Therefore, in a herd of 1000 sows, 500 new gilts are introduced every year (~48 every 5 weeks), and 450 piglets are born every week replacing one third of the piglet population. We speculate that the high turnover rate of piglets and the continuous introduction of gilts into breeding herds are significant factors associated with IAV diversity, the introduction of novel IAVs into pig breeding herds, and the emergence of reassortant IAVs.

The co-circulation of two or more IAVs allows genetic reassortment and the emergence of novel IAVs (23, 144). In North America, the genetic diversity of swine IAVs has increased dramatically since the emergence of the triple reassortant internal gene (TRIG) cassette in 1998 (16, 75) and the introduction of several human IAVs into swine populations, including the 2009 pandemic virus (15, 61). In this study we found several reassortant viruses, which included reassortment events with viruses isolated in this study (suggestive of reassortment happening during the study period), and reassortment events with viruses not isolated in this study. Whether the reassorted viruses identified in this study emerged within the herds studied, in the herds that supplied the gilts, or in other pig herds is not clear. However, in this study the isolation of three IAVs with more than 8 gene segments (mixed genotypes) indicated that the conditions for reassortment to occur existed during the study period. In contrast, the identification of gene segments that were closer to other IAVs not isolated in this study, but currently circulating in the USA, suggest that external sources (e.g. other pigs) are also important for the emergence of reassortant viruses in breeding herds. Therefore, we believe that in pig breeding herds, the interventions to control IAV infection should not only target the transmission within herds, but also minimize the risk of new IAV introductions.

The ability of IAVs to exchange gene segments over time not only increases the mechanisms of virus diversification but also may allow genetic traits (e.g signature mutations) to move between IAVs. Seventy-nine percent of IAVs

isolated in our study contained a signature mutation (S31N) in the matrix gene (segment 7) that might confer resistance to amantadine (138). This high prevalence of amantadine resistance is in agreement with two previous studies on swine IAVs (145, 146). Interestingly, amantadine is not FDA labeled for use in pigs in the USA. Moreover, the incidence of human IAVs resistant to adamantanes changed from 0.4% in 1994 (147) to 15.5% in 2006 (148) and could be associated with antiviral usage in humans. Whether the high frequency of S31N in the matrix gene among swine IAVs is due to a random events, as indicated by Baranovich et al., 2015 (145), or due to reassortment events, as indicated by Krumbholz et al., 2009 (146) is not clear and must be further investigated. Nevertheless, multiple introductions of human IAVs into swine populations have led to the establishment of certain IAV gene lineages (61) and we speculate that in that process, genetic resistance signatures to antivirals, such as amantadine, have also been incorporated. Hence, these introductions and establishment of “foreigner” IAV gene segments into swine populations could likely result in unique, for example drug-resistant, genotypes.

Our results do not represent the overall dynamic of IAV infections in pig breeding herds in the Midwestern USA given our herd-selection bias. Additionally, we could have missed some IAVs over time given our study design and prior genotype selection by culturing the viruses rather than sequencing it directly from the swabs. Our sample size did not allow us to explore further the association between the genetic diversity of IAVs and pig sub-populations (e.g. we cannot

test which subpopulation had more or less IAV diversity). Moreover, we could not identify the source or directionality of transmission of the IAVs in these herds. Given that new gilts are introduced in to the farm from external sources, we believe that they are the most likely source of the introduction of genetically distinct IAVs, although air (20) and fomites (19) can also represent a risk for new IAV infections in pig herds. Alternatively, gilts could have also become infected with resident viruses after arrival. Future studies sampling gilts at arrival to breeding herds could clarify the relative importance of gilts as a source of new IAVs to the breeding herd versus their role in amplifying resident viruses. Moreover, caution should be taken when using our estimates for the linear association between the distance-to-the-root of the phylogenetic tree and time (in days) for each genetic lineage analyzed in this study. The observations used (sequences) are not independent from each other because the virus is evolving from a common ancestor (149) hence, the assumption of independence is violated. We acknowledge that there are other methods (e.g Bayesian analysis) to estimate evolutionary rates and TMRCA. Nevertheless, our results are comparable to studies using Bayesian analysis to study swine IAVs (61) and FastTree2 has been confirmed as a robust method for phylogenetic studies of large datasets (135).

In conclusion, our study demonstrates the complex and dynamic occurrence and maintenance of IAV infections in pig breeding herds. Complete genome amplification and NGS technologies allowed us to characterize with more

precision the complete genome of IAVs over time. We showed that IAVs can be sustained for prolonged periods of time and that distinct IAVs can co-exist within and between subpopulations in these herds. Thus, pigs in breeding herds are repeatedly exposed to distinct IAVs over time. Our results also indicated that pig population dynamics and not only the viral mechanisms of genetic diversification should be taken into account in elucidating the diversity and evolution of swine IAVs. We speculate that if transmission of IAVs is reduced in the breeding herds the distribution of IAVs to other pig sites after weaning will be minimized. Hence, understanding the epidemiology and evolution of IAVs in pig breeding herds will allow us to design more effective strategies to reduce the impact of IAV infections on swine health and to minimize the risk to public health.

Table 6. Total number of nasal swabs collected, samples positive to influenza A virus (IAV) by reverse real time polymerase chain reaction (RRT- PCR), and IAV isolates distributed by farm, subpopulation and subtype.

Farm	Subpopulation	Total number of swabs collected	Number of IAV RRT-PCR positive samples used for virus isolation[+] (%)	Number of IAV isolates (%)	IAV Subtypes
1	New gilts	119	1 (0.8)	0 (0)	NA
	Gilts	360	0 (0)	NA	NA
	Piglets	358	21 (5.9)	12 (57.1)[++]	H1N1 (n=4), H3N2 (n=7)
	Subtotal	837	22 (2.6)	12 (54.5)[++]	H1N1 (n=4), H3N2 (n=7)
2	New gilts	119	0 (0)	NA	NA
	Gilts	360	0 (0)	NA	NA
	Piglets	358	7 (2.0)	5 (71.4)	H1N1 (n=2), H1N2 (n=2), H1N2/H3N2 (n=1)
	Subtotal	837	7 (0.8)	5 (71.4)	H1N1 (n=2), H1N2 (n=2), H1N2/H3N2 (n=1)
3	New gilts	268	72 (26.9)	40 (55.6)	H1N1 (n=12), H3N2 (n=28)
	Gilts	359	19 (5.3)	14 (73.7)	H1N1 (n=2), H3N2 (n=10), H1N1/H3N2 (n=1), H1N1/N2 (n=1)
	Piglets	360	20 (5.6)	20 (100)	H1N1 (n=3), H1N2 (n=15), H3N2 (n=2)
	Subtotal	987	111 (11.2)	74 (66.7)	H1N1 (n=17), H1N2 (n=15), H3N2 (n=40), H1N1/H3N2 (n=1), H1N1/N2 (n=1)
4	New gilts	60	0 (0)	NA	NA
	Gilts	349	19 (5.4)	4 (21.1)	H1N2 (n=4)
	Piglets	360	21 (5.8)	15 (71.4)	H1N1 (n=7), H1N2 (n=5), H3N2 (n=3)
	Subtotal	769	40 (5.2)	19 (47.5)	H1N1 (n=7), H1N2 (n=9), H3N2 (n=3)
5	New gilts	60	0 (0)	NA	NA
	Gilts	340	0 (0)	NA	NA
	Piglets	360	27 (7.5)	14 (51.9)	H1N1 (n=1), H3N2 (n=13)
	Subtotal	760	27 (3.6)	14 (51.9)	H1N1 (n=1), H3N2 (n=13)
Total		4190	207 (4.9)	124 (59.9)	H1N1 (n=31), H1N2 (n= 26), H3N2 (n=63), H1N1/H3N2 (n=1), H1N1/N2 (n=1) H1N2/H3N2 (n=1)

[+] Only positive samples with a cycle threshold ≤ 35 were used for IAV isolation.

[++] One IAV isolate from piglets was not successfully sequenced

NA: Not applicable

Table 7. Simple linear regression estimates for the association between the distance-to-the-root of the phylogenetic tree for each gene segment dataset and time (in days).

Thirteen different genetic lineages of swine influenza A virus gene segments (datasets) were used. January 1st 2003 was set as day 1 and the best fitting root for each tree was estimated using Path-O-Gen v1.4.

Segment	Sequences from this study	IRD & USDA[+] sequences	Total sequences	Intercept (α)	(β1)[++]	95% Confidence interval		Adjusted R2	TMRC[+++] (days)	Date (TMRCA)
						2.5%	97.5%			
Seg 1 (PB2)	125	1404	1529	1.51E-02	8.64E-06	8.46E-06	8.82E-06	0.86	-1747	3/21/98
Seg 2 (PB1)	124	1274	1398	1.66E-02	8.86E-06	8.66E-06	9.06E-06	0.85	-1873	11/15/97
Seg 3 (PA)	125	1387	1512	2.00E-02	7.07E-06	6.90E-06	7.24E-06	0.82	-2823	4/10/95
Seg 4 (H1 classic)	33	1009	1042	5.62E-02	1.21E-05	1.18E-05	1.24E-05	0.83	-4650	4/9/90
Seg 4 (H1 delta)	25	482	507	-3.47E-02	1.36E-05	1.30E-05	1.42E-05	0.80	2560	1/3/10
Seg 4 (H3)	67	771	838	4.12E-02	1.32E-05	1.28E-05	1.36E-05	0.82	-3122	6/14/94
Seg 5 (NP)	125	1554	1679	1.65E-02	7.79E-06	7.53E-06	8.06E-06	0.66	-2121	3/12/97
Seg 6 (N1)	33	690	723	2.86E-02	9.34E-06	9.04E-06	9.65E-06	0.83	-3057	8/19/94
Seg 6 (N2)	93	1729	1822	2.54E-02	1.51E-05	1.46E-05	1.55E-05	0.71	-1688	5/19/98
Seg 7 (M pandemic)	86	1557	1643	-1.62E-02	7.82E-06	7.52E-06	8.12E-06	0.61	2078	9/8/08
Seg 7 (M non pandemic)	38	959	997	1.73E-02	5.55E-06	5.31E-06	5.79E-06	0.67	-3125	6/11/94
Seg 8 (NS pandemic)	38	223	261	-1.97E-02	9.12E-06	8.23E-06	1.00E-05	0.61	2156	11/26/08
Seg 8 (NS non pandemic)	88	1362	1450	1.61E-02	4.64E-06	4.40E-06	4.87E-06	0.51	-3465	7/7/93

[+] Influenza Research Database (IRD) and United States Department of Agriculture (USDA).
[++] β1 is expressed as number of nucleotide difference between sequences for each day increment over time between January 1st 2003 and October 16th 2014.
[+++] Estimated time to most recent common ancestor (TMRCA).

Figure 3. Phylogenetic relationships between antigenic gene segments of influenza A virus (IAV) isolates from this study and IAVs circulating in the USA between January 1st 2003 and October 16th 2014.

Panels a, b and c represent hemagglutinin (HA) sequences and panels d and e represent neuraminidase (NA) sequences. The HA sequences under analysis clustered into 7 distinct study clades (SC): 1 (blue); 2 (purple); 3 (pink); 4 (lime); 5 (green); 6 (turquoise); and 7 (aqua). Neuraminidase (NA) sequences are colored according to HA study clades and the asterisk "*" indicates one N2 sequence that did not cluster within the expected clade.

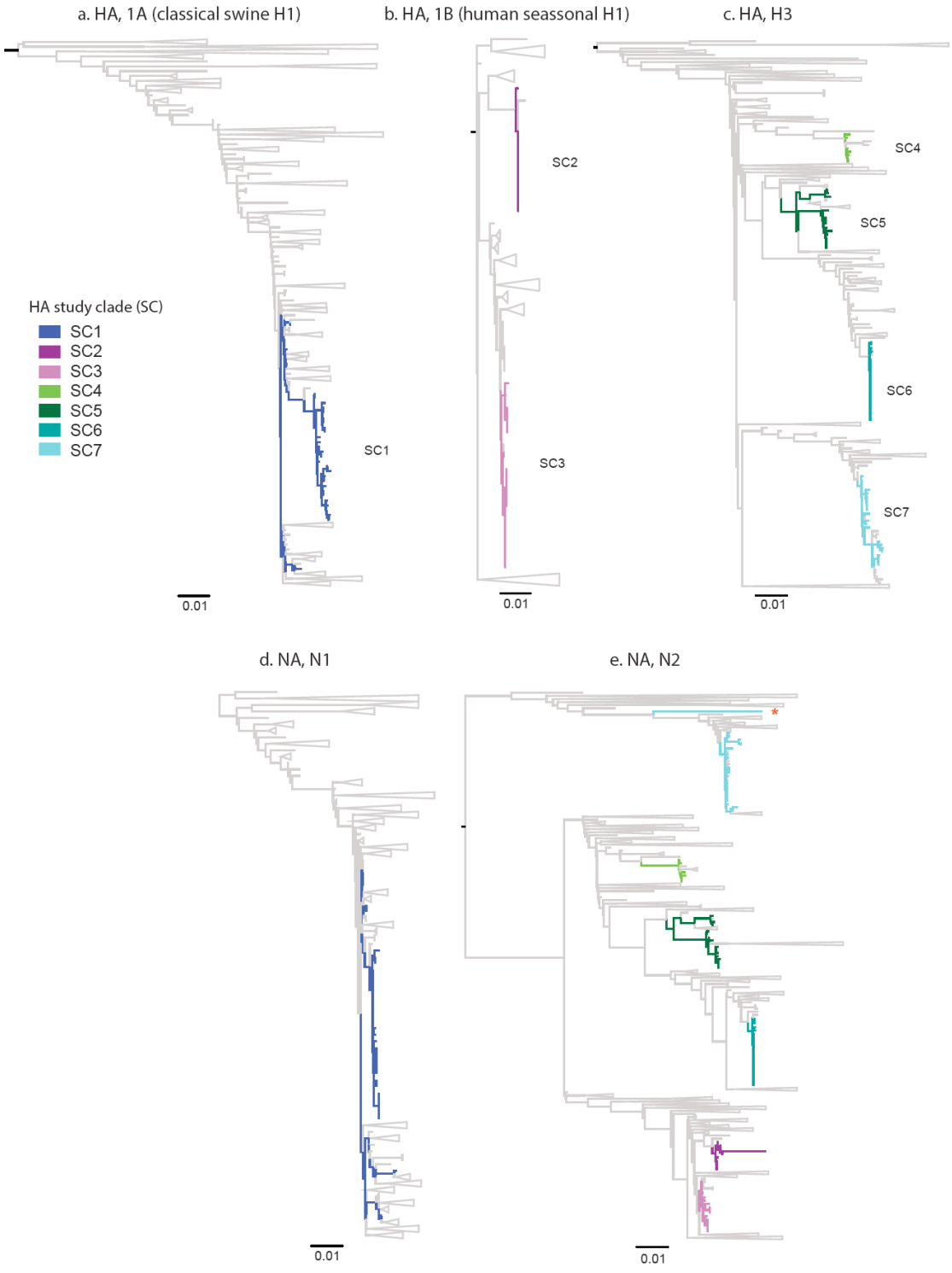
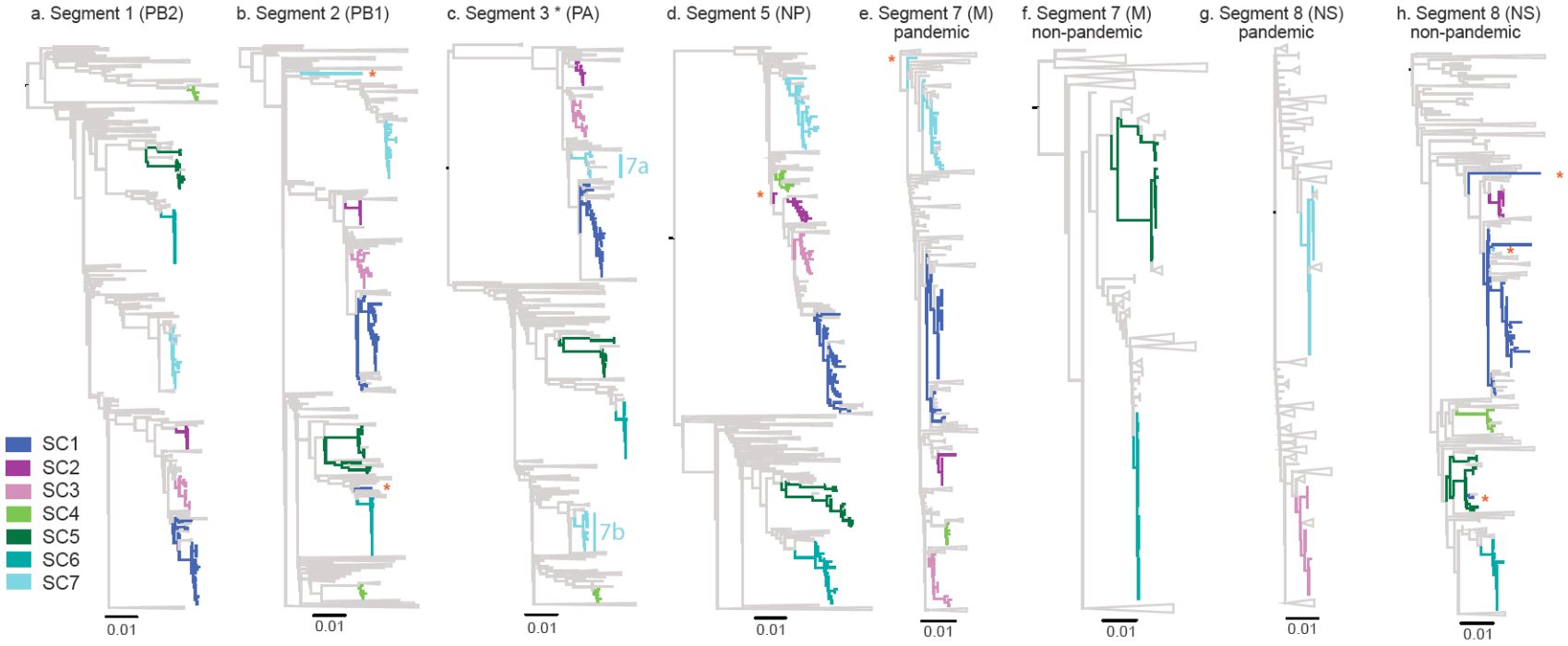

a. HA, 1A (classical swine H1)
b. HA, 1B (human seassonal H1)
c. HA, H3
HA study clade (SC)
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC1
SC2
SC3
SC4
SC5
SC6
SC7
0.01
0.01
0.01
d. NA, N1
e. NA, N2
0.01
0.01

Figure 4. Phylogenetic relationships of the internal gene segments of influenza A virus (IAV) isolates from this study and IAVs circulating in the USA between January 1st 2003 and October 16th 2014.

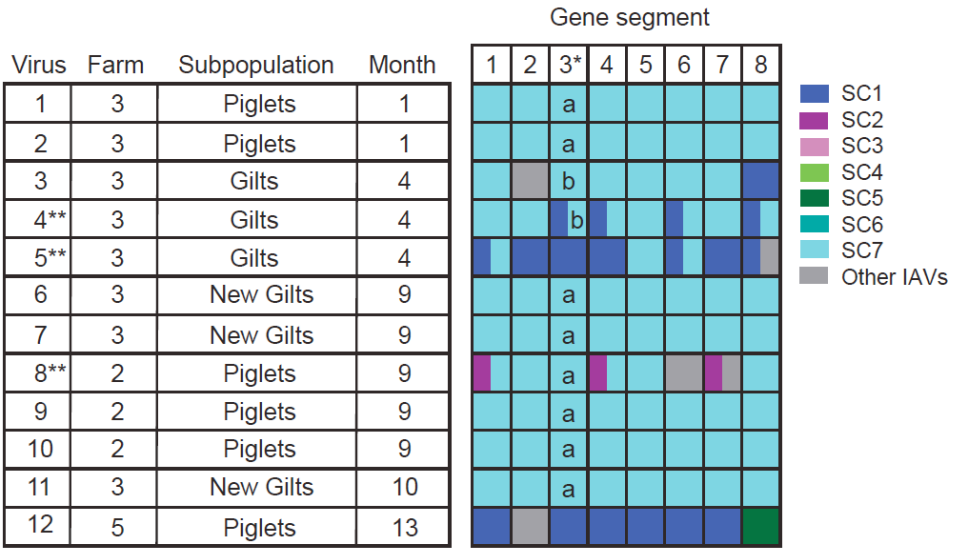
Tree leaves representing sequences from this study are colored according to the hemagglutinin (HA) study clades (SC1 – SC7) illustrated in Fig 3 and asterisks "*" indicate sequences that did not cluster within the expected clade.



* 7a and 7b indicate two distinct PA clades found for viruses containing HA within SC7.

Figure 5. Reassortant influenza A viruses (IAVs) found during the study period.

The genome constellations for 12 IAVs are shown distributed by farm, subpopulation and month. Gene segments are color-coded based on HA study clades: 1 (blue), 2 (purple), 3 (pink), 4 (lime), 5 (green), 6 (turquoise), and 7 (aqua). Gray indicates sequences that were closer to other IAVs not isolated in this study.



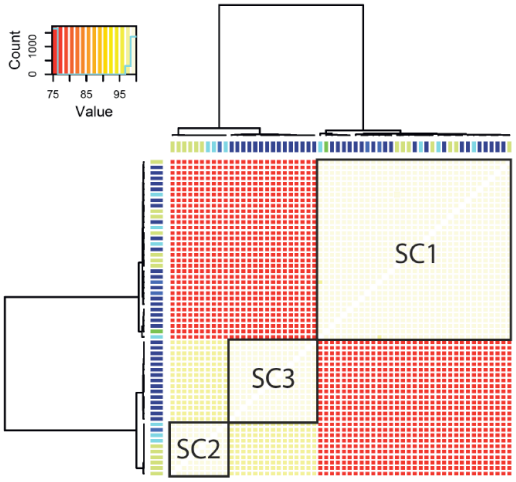
* Viruses with HA genes clustering within SC7 (aqua) had two different PA genetic lineages (a and b) as illustrated in Fig 2c

** IAVs with more than one antigenic subtype (mixed viruses)

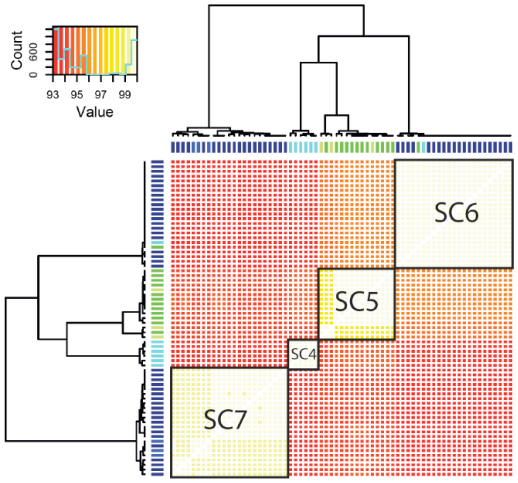
Figure 6. Hemagglutinin and neuraminidase sequence comparison within and between farms.

Dendrograms and heat maps illustrate the pairwise identity matrix (ClustalX alignment) for hemagglutinin (subtypes H1 and H3) and neuraminidase (subtypes N1 and N2) sequences. The color key and histogram for the pairwise identity comparison is indicated top left side in each panel. The side color bar in each dendrogram indicates the farm number for each sequence as follows, farm 1 (aqua), 2 (purple), 3 (blue), 4 (olive), and 5 (green). Additionally, the distribution of HA sequences by study clades 1 to 7 (SC1-SC7) is indicated in panels a and b.

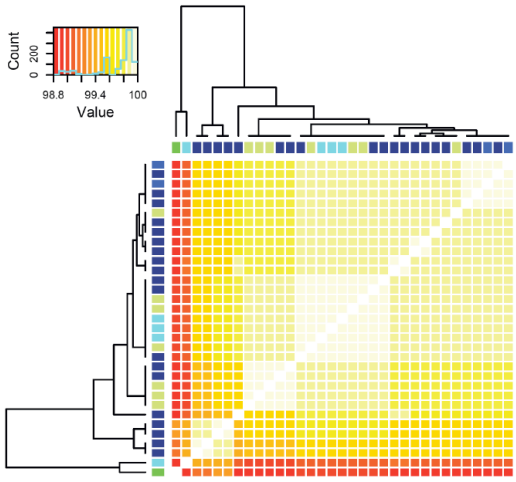
a. Hemagglutinin (H1), n=58



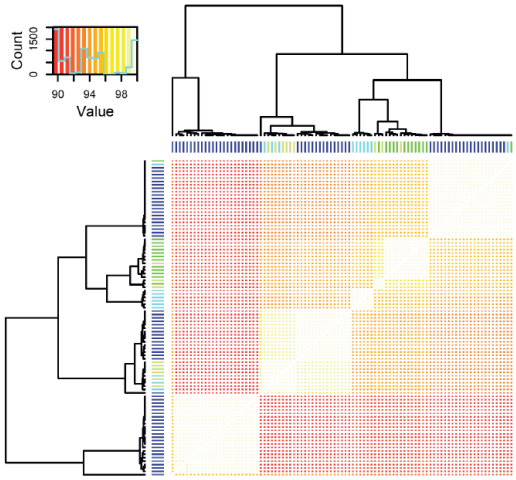
b. Hemagglutinin (H3), n=67



c. Neuraminidase (N1), n=33



d. Neuraminidase (N2), n=93



Farm 1 Farm 2 Farm 3 Farm 4 Farm 5

Figure 7. Influenza A virus (IAV) isolates distributed by hemagglutinin (HA) study clades (SC1-SC7), farm (F1-F5), subpopulation (new gilts (NG), gilts (GL) and piglets (PG)) and month.

Each column represents a sampling month and rows indicate farm and pig subpopulation. Within farm, each box represents a sampling event and missing boxes indicate months when samples were not collected. White boxes indicate sampling events when IAV was not isolated. Bars within boxes indicate IAV isolation and are color-coded according to HA phylogenetic study clades SC1 to SC7 illustrated in Fig 3.

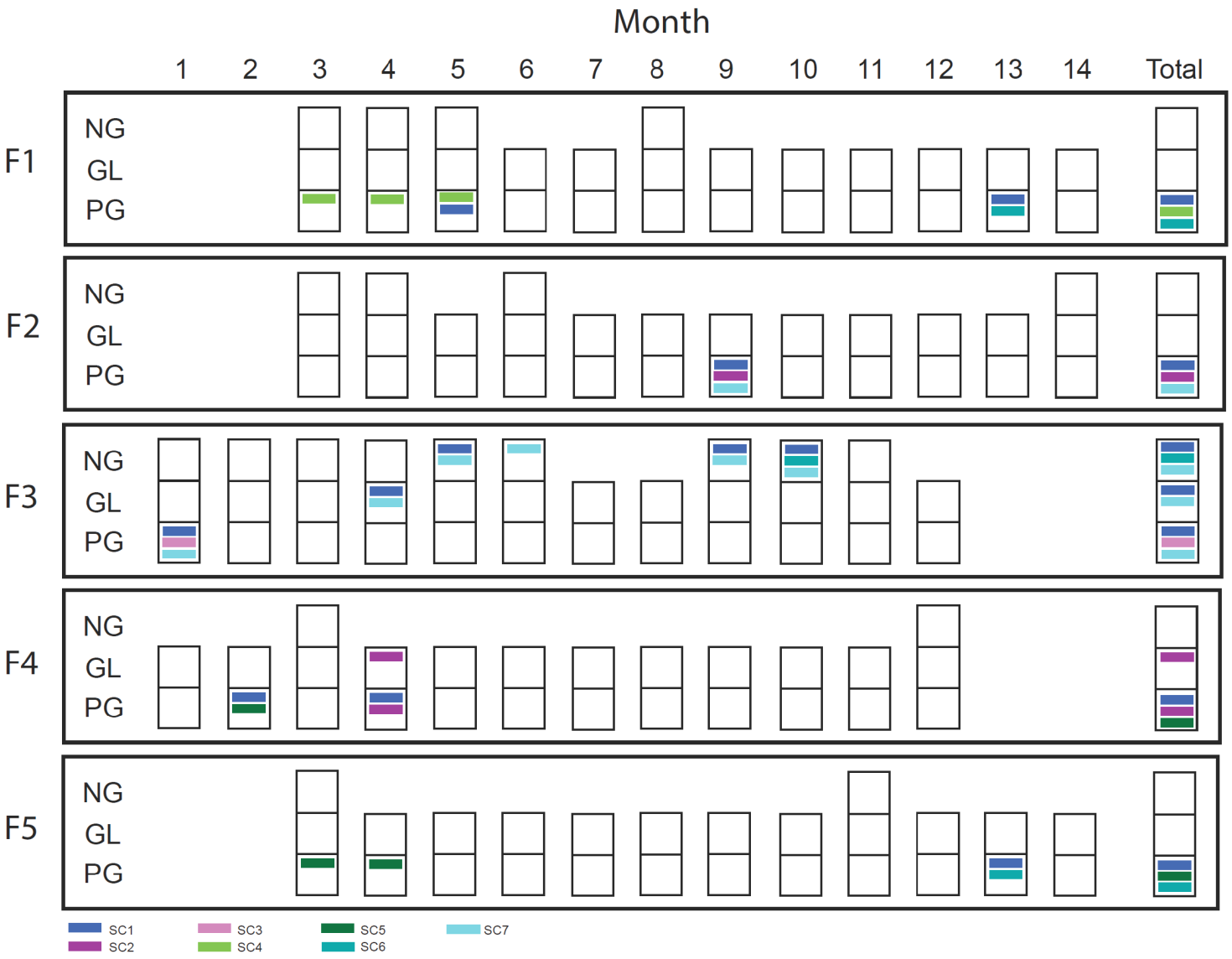
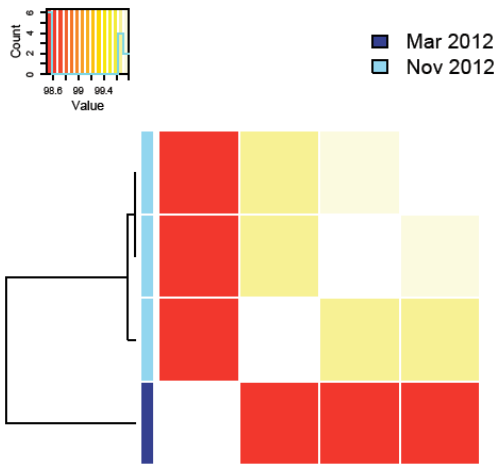

Month
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Total
F1
F2
F3
F4
F5
NG
GL
PG
SC1
SC2
SC3
SC4
SC5
SC6
SC7

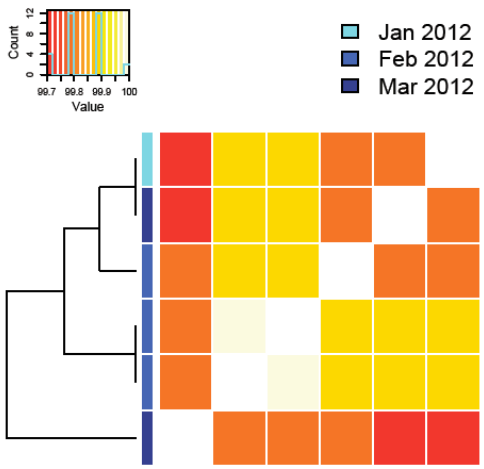
Figure 8. Dendrograms and heat maps illustrating the pairwise identity matrix (ClustalX alignment) between hemagglutinin sequences of the same study clade (SC) identified within the same farm over time.

a) Farm 1 subtype H1 SC1, b) Farm 1 subtype H3 SC4, c) Farm 3 subtype H1 SC1, d) Farm 3 subtype H3 SC7, e) Farm 4 subtype H1 SC1, f) Farm 5 subtype H3 SC5. The side color bar in each dendrogram indicates the month when each virus was collected.

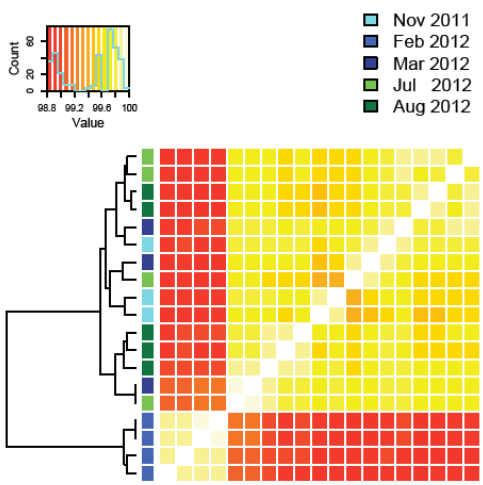
a. Farm 1, Subtype H1 SC1



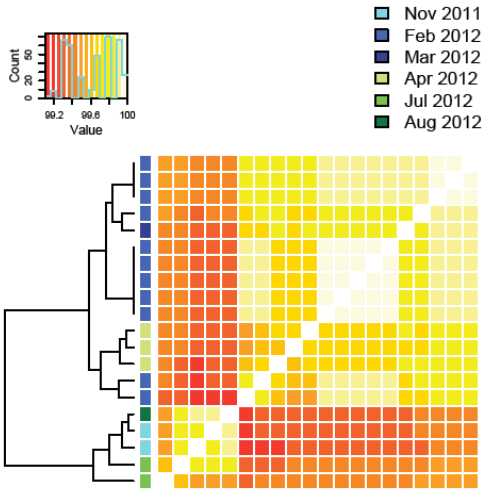
b. Farm 1, Subtype H3 SC4



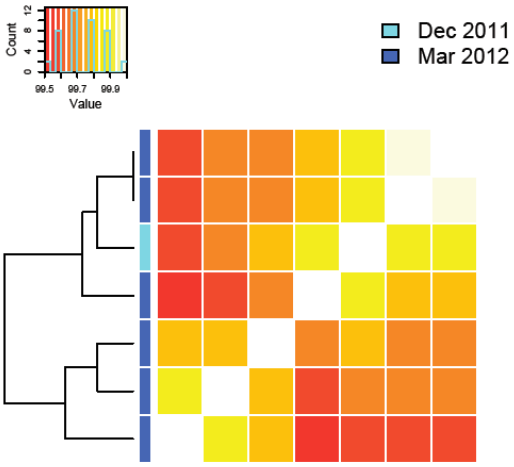
c. Farm 3, Subtype H1 SC1



d. Farm 3, Subtype H3 SC7



e. Farm 4, Subtype H1 SC1



f. Farm 5, Subtype H3 SC5

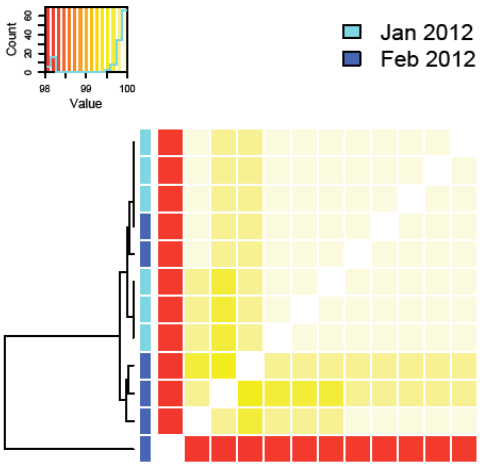
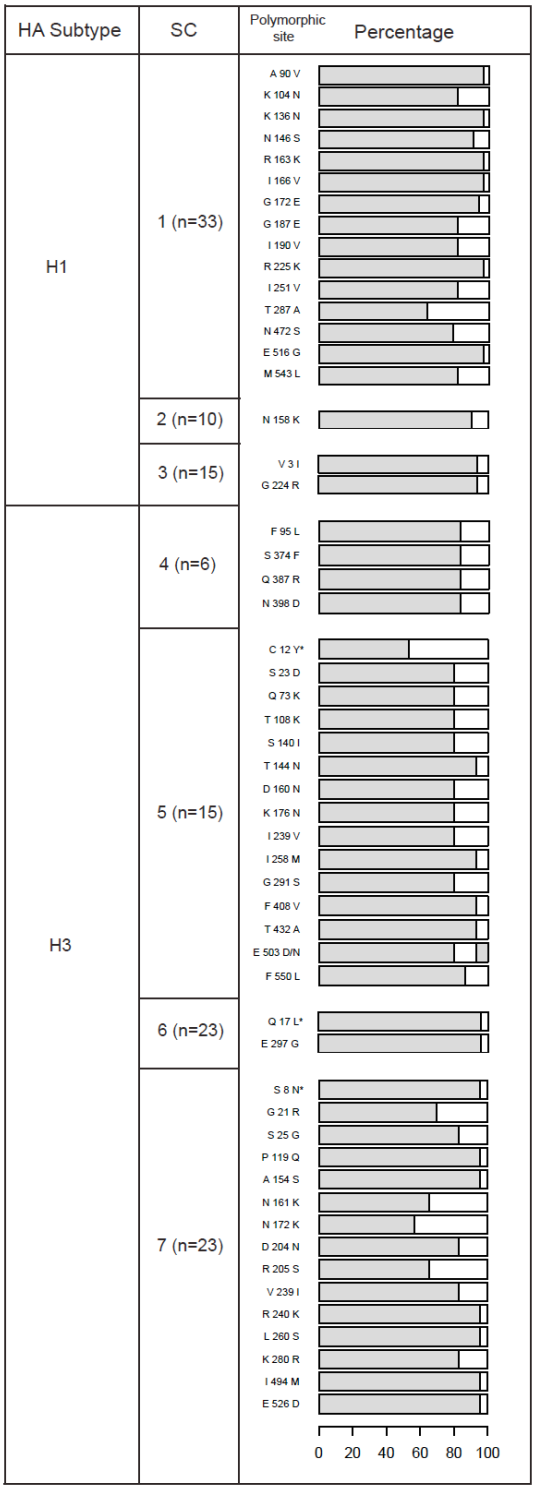


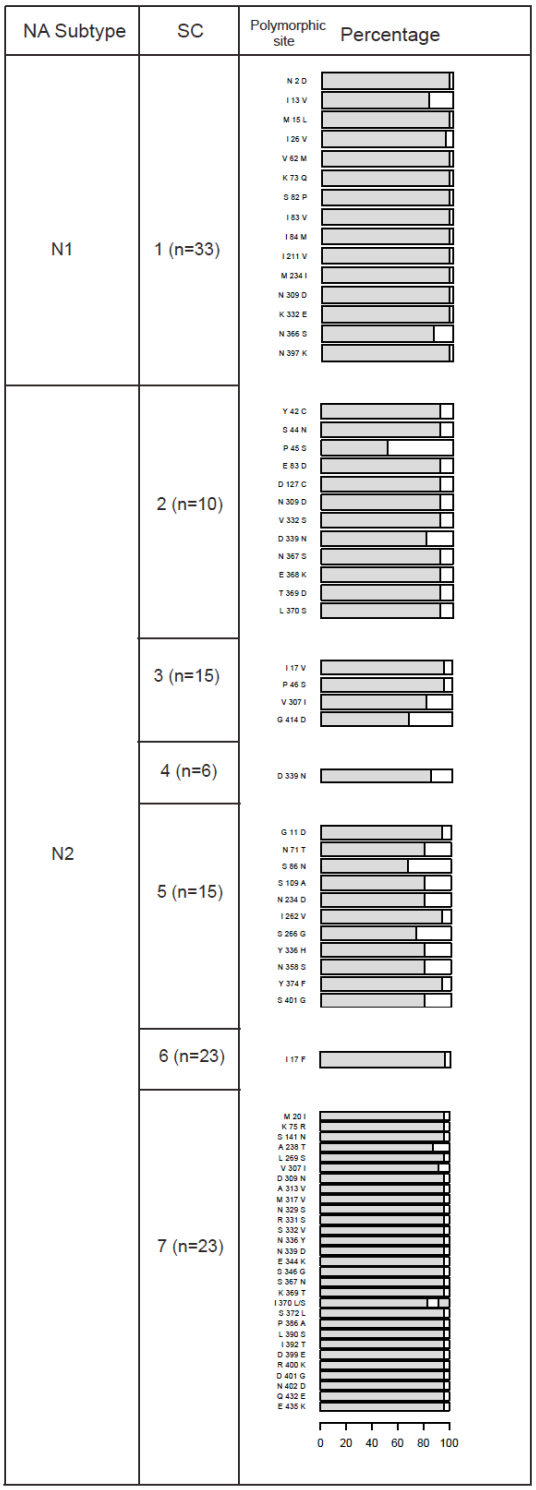
Figure 9. Percentage of polymorphic amino acid sites in the hemagglutinin (panel a) and neuraminidase (panel b) of influenza A virus (IAV) isolates.

IAV subtype, HA study clade, and number of sequences used in each comparison is indicated. The polymorphic site indicates the most frequent amino acid found (first letter), the position (number) and the variant amino acid (second letter). Horizontal bars represent the percentage distribution of each amino acid at each polymorphic site. (*) Indicates polymorphic sites within the signal peptide of the HA.

a. Hemmaglutinin



b. Neuraminidase



Chapter 4: Antigenic drift of H1N1 influenza A virus in pigs with and without passive immunity

This work has been published in:
Diaz A. Allerson M. Culhane M. Sreevatsan S. Torremorell M. (2013)
Influenza and Other Respiratory Viruses 7 (Suppl. 4), 52–60.

Introduction

Influenza A viruses (IAV) carry a segmented negative sense RNA genome that is able to reassort with other IAV strains (antigenic shift) and/or evolve by the accumulation of mutations throughout the genome (antigenic drift) (9). Two proteins are known as major IAV antigens, hemagglutinin (HA) and neuraminidase (NA), and their genotypes are associated to the host species that each virus infects (41). Accumulation of mutations and gene exchange between IAV during infection is expected to influence viral fitness within and transmission between species (43, 150). Factors responsible for mutations in IAV are not completely understood; however, a viral non-proofreading polymerase (48), immune selection (151), and intra-host characteristics are known to play key roles (5, 73, 126). Although wild waterfowl are considered the natural reservoirs for IAV, pigs can be intermediate hosts(49) with a propensity for generating reassortant viruses(13) and sustaining infections that result in new viruses of risk to other species, including humans (152).

Herd prevalence estimates indicate that IAV infections are endemic and widespread in pigs (120). In the US, after the emergence of the 2009 pandemic H1N1 (pH1N1) in swine, the HA gene of H1 subtype clustered in 5 different phylogenetic groups (α, β, γ, δ1, and δ2) which illustrates the broad diversity of IAV in pigs (56). Active surveillance in U.S. swine has shown that 90% of the herds surveyed throughout a two-year period tested virus positive for IAV (17). Furthermore, these herds were positive for multiple IAV strains of a variety of subtypes resulting from several reassortment events (23). Such IAV diversity in pigs increases the challenges faced in both understanding IAV evolution and controlling IAV in pigs.

Suckling pigs serve as an important source of IAV since they can be asymptomatically infected and can transport the virus to multiple geographical locations at weaning. Virus shedding occurs even if passive (38) or active immunity (64) are present, and different genotypes of IAV can co-circulate in swine populations regardless of their immune status (73) which may result in the emergence of novel reassortant strains. However, little is known about genetic diversity and selective evolution of IAV in suckling piglets with passive immunity. Given the central role that suckling pigs play in the dissemination and emergence of IAV strains in pig farms, the objective of this study was to evaluate the degree of antigenic drift in the HA of a H1N1 virus in pigs with passive immunity under experimental conditions.

Materials and Methods

Study design and sample selection:

An IAV negative swine herd was selected for the study. The herd was considered negative for IAV based on serology (NP ELISA, IDEXX, Ref) and clinical history. Sows were either vaccinated with an experimental vaccine (PASSIV-VAC) prepared with H1N1 influenza virus A/Swine/IL/02450/08 (α cluster, NCBI accession number: CY099052.1) or left unvaccinated (NAIVE). The vaccine was an adjuvant inactivated vaccine and administered intramuscularly at 5 and 2 weeks prior to farrowing.

At weaning, 30 and 39 piglets were selected from the PASSIV-VAC and NAIVE groups respectively and moved to the University of Minnesota animal isolation units. Nine pigs from the NAIVE group were randomly selected to be used as seeder pigs, and were infected intra-nasally and intra-tracheally with 0.5 ml of an inoculum containing 10-7 TCID50 of influenza A/Swine/IA/00239/04 H1N1 virus

(β cluster, NCBI accession number: EU139832.1). The remaining pigs were randomly allocated into groups of 10 pigs (3 replicates per group).

Sows in the PASSIV-VAC group were confirmed not to have been exposed to IAV before vaccination by testing serum samples by ELISA assay (FlockChek® Avian Influenza MultiS-Screen Antibody Test Kit, IDEXX Laboratories Inc., Westbrook, ME, USA). Piglets in the NAIVE group were also confirmed to not have immunity to IAV via serum sample tests using the same ELISA assay. Additionally all piglets in the PASSIV-VAC and NAIVE were confirmed to be IAV negative prior to exposure to the seeder pig by testing nasal swabs by RRT-PCR (85).

After pigs were exposed to at least one seeder pig, nasal swabs were collected daily for 14 days and tested by RRT-PCR targeting the matrix gene (M)(85). The seeder pigs remained with the rest of the pigs throughout the study. For sequencing purposes, one positive sample from each animal was conveniently selected based on the lowest cycle threshold (CT) value obtained on the matrix RRT-PCR reaction and all infection days were represented. Full HA gene amplification was performed directly from the selected nasal swabs. Blood samples were collected prior to infection and at the termination of the study. All sera were assayed by hemmaglutination inhibition (HI) test (90) against the challenge virus, A/Swine/IA/00239/04, and the vaccine virus, A/Swine/IL/02450/08.

Sequencing and sequence analysis:

In order to establish virus HA genetic and protein relatedness between the challenge and vaccine virus, full length HA sequences of A/Swine/IA/00239/04 (EU139832.1) and A/Swine/IL/02450/08 (CY099052.1) were first aligned and compared using Clustal W (MegAlign, LaserGene Core 9. DNASTAR).

From each selected sample, HA gene was amplified using previously described primers (93), and PCR products confirmed by electrophoresis in 2% agarose gel and stained with ethidium bromide. The expected band for HA (1780 base pairs) was excised from the gel and eluted using QIAquick Gel Extraction Kit (QIAGEN) following manufacturers' recommendations, and submitted to the Biomedical Genome Center (BMGC) of the University of Minnesota for sequencing using a primer walking scheme.

The trace files obtained from each sample were assembled using the HA sequence of A/swine/New Jersey/11/76(H1N1) as a reference. The template was then removed and the consensus sequence for each sample established. All contigs were initially assembled using the default parameters of SeqManPro (LaserGene Core 9, DNA-STAR) and the quality of each nucleotide trace was evaluated visually in each position. Only nucleotide calls with clear peaks were considered for analysis. Each trace file was trimmed when peaks were considered ambiguous to obtain complete sequences with good quality reads. Each consensus sequence was annotated using the influenza annotation tool of the NCBI (http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/annotation.cgi), and trimmed accordingly to obtain the HA coding region.

Complete sequences within each group were aligned and compared by Clustal W (MegAlign, LaserGene Core 9. DNA-STAR) using the consensus from all sequences as reference, nucleotide differences were classified as either transitions (purine↔purine or pyrimidine↔pyrimidine) or transversions (purine↔pyrimidine) according to the base pair interchange identified, and their phylogenetic distance was compared using a median-joining network algorithm (153). The proportion of synonymous (SM) and non-synonymous (NSM) mutations were assessed and compared between groups. In some cases, full-length HA sequences were not obtained, thus short reads for those samples were aligned and polymorphisms were inferred for only the regions sequenced.

Statistical Methods:

The nucleotide substitution mean was obtained for each group (sum of the number of nucleotides substituted/total number of nucleotides sequenced in all samples) and compared between groups using a one-tail t-test. Results were considered statistically significant at $p<0.05$. Additionally the odds of non-synonymous mutations (NSM/SM) were compared between groups using only complete HA sequences. Finally the frequency of non-synonymous mutations present more than once was compared between groups (including complete and incomplete sequences). For these last two comparisons tabular methods were used and considered statistically different when $\chi 2$ $p < 0.05$. If the count per cell in each comparison was lower than 5 then a Fisher-exact test was used.

HA protein models:

To better visualize the location of the amino acid differences identified, the structure of the translated HA1 regions were modeled using A/Swine/Iowa/15/30 (H1N1) as template (PDB:1RUY) with the tools offered in swissmodel.expasy.org (154). Each amino acid substitution detected was mapped within the predicted model for the reference sequence and its location compared with the five antigenic sites previously identified for A/PR/8/1934 (PDB: 1RU7) (155, 156) using PyMOL™ 2010. The nucleotide numbering for HA followed the numbering for the entire HA gene, while the amino acid numbering was segregated by sig_peptide, HA1 and HA2 portions, such that each region of the molecule was identified by a specific amino acid where they began and/or ended.

Results

The number of days that pigs tested positive, and the PCR positive samples selected for sequencing and comparison are shown in Table 8. Prior to exposure

all pigs in the NAIVE group tested negative to IAV by ELISA (S/N ratio ≥0.673) and all pigs in the PASSIV-VAC group tested positive (mean: 0.254, min: 0.09, max: 0.495). HI titers for each group before and after contact to the seeder pig are shown in Table 9. HI titers against the challenge strain increased in both groups after infection.

HA gene sequence of A/Swine/IA/00239/04 and A/Swine/IL/02450/08 aligned within the β and α H1 clusters of swine IAV, respectively, and both followed the same reading frame described for the human pandemic influenza virus A/California/04/2009(H1N1) (GenBank: FJ966082.1). The overall nucleotide and amino acid identity for the HA gene and hypothetical proteins, as well as the identity for each expected antigenic site are summarized in Table 10.

Both SM and NSM were observed in both groups of pigs (Table 11). Differences were observed within sig-peptide and HA1 region for the NAIVE group (16 complete sequences), and within HA1 and HA2 for the PASSIV-VAC group (9 complete sequences). Additionally 12 and 16 partial sequences were obtained in the PASSIV-VAC and NAIVE groups respectively, and their polymorphisms are summarized in Table 12. The total number of polymorphisms identified in both groups led to eight different HA alleles (complete sequences different one from each other), and their phylogenetic relatedness is illustrated in Fig 10.

The nucleotide substitution mean for the NAIVE (2.94×10^{-4}) was lower ($p=0.042$) when compared to the mean in the PASSIV-VAC group ($6.4 \times ^{10-4}$). The odds of NSM in the PASSIV-VAC group was 1.29 times higher compared to the NAIVE group (Table 13), however this difference was not statistically significant ($p>0.05$). Overall, the most common non-synonymous mutation in both groups was at position 515 and the odds of this mutation (G515A) in the PASSIV-VAC group was 1.6 times higher compared to the NAIVE group (Table 14), however this was not statistically different ($p=0.38$). Table 15 summarizes the hypothetical amino acid substitutions found in all groups and Fig 11 shows

their location within the HA1 protein model. All changes in the NAIVE group were mapped within antigenic site B, while changes in the PASSIV-VAC group were located within antigenic site A, B, and D. The substitution identified at position 155 found in both groups, indicate a change from a polar to a charged amino acid, all the other substitutions found in the PASSIV-VAC group did not change the amino chemical nature, but the substitution found in the NAIVE at position 189 did (polar to hydrophobic).

Discussion

IAV in pigs represents a constant risk to other species including humans since pigs serve as a reservoir for IAV that can result in zoonotic infections of pandemic proportions (10). Although it is known that factors such as active immunity can play a role at inducing IAV change, little is known about the degree of virus diversity found in pigs with passive immunity. In this study, we evaluated genetic and antigenic changes in the HA of an H1N1 IAV in pigs with and without passive immunity under experimental conditions with emphasis on changes observed in young pigs early in the infection process. Although nucleotide substitution in swine IAV is believed to be lower compared to human viruses (60), this study demonstrates that genetic changes can occur in young pigs early during infection and those changes can induce amino acid changes located within antigenic sites. Furthermore, both SM and NSM were observed in the HA protein of pigs with and without passive immunity, and sequences from both groups are genetically related regardless of immune status. Even though the nucleotide substitution mean was higher in the PASSIV-VAC group compared to the NAIVE group, the odds of NSM in the PASSIV-VAC group was not significantly higher compared to the NAIVE group. Changes in the chemical nature, especially of G155E identified in 11 pigs (5 and 6 in the NAIVE and PASSIV-VAC respectively), indicate that there might be specific changes that occur in early infection to improve viral fitness in a new host. However, the biological significance of this type of change as it relates to immune pressure

needs to be further investigated. The dynamic distribution of IAV subtypes in pigs, and the ability of the virus to change early in infection has been described and suggests that allele fixation can occur rapidly (73).

Immune pressure and vaccination have been associated with virus change in other animal models (63). In swine, the use of influenza vaccines is common and usually vaccines are administered to breeding females prior to farrowing. According to USDA National Animal Health Monitoring System (NAHMS) data, 70% of large breeding herds (> 500 sows) in the US vaccinate against IAV (http://www.aphis.usda.gov/animal_health/nahms/swine/downloads/swine2006/Swine2006_is_vacc.pdf). Pre-farrowing vaccination is practiced to enable transfer of passive immunity and minimize clinical disease to progeny. However, passive immunity is generally insufficient to prevent transmission of IAV in pigs (38). It is logical that immune pressure created by vaccination of pigs would drive virus change. However, in our study, passive immunity did not affect virus transmission and we did not see significant genetic differences between immune and naïve pigs. The similarity of virus change between groups in this study could also be a result of the phenotypic differences between the virus used to prepare the vaccine and the virus used to infect the seeder pig. However these results are in agreement with a recent study where it was shown that vaccination did not appear to have a major effect on the genetic structure of intra-host viral populations through immune selection (73).

We selected samples based on the amount of genetic material present in the sample and ensured that there was representation of infection through all days of the study. Most of the mutations were observed one time post infection. Only one mutation was seen in multiple animals. This mutation was at position 155 of HA1 and it was observed in naïve and immune pigs. This mutation resulted in glycine to glutamate substitution, which changes the predicted conformation of the HA structure. Substitution at this same site was also described by Hensley et al., 2009, but only in immune mice after serial passages. Furthermore the mutation,

E156K of (63) A/Puerto Rico/8/34(H1N1), which aligns with 155 of A/Swine/IA/00239/04 in this study, was mapped within HA antigenic site B and was associated to immune escape. Additionally a mutation at position 155 was also described by Kuroda et al., 2010 (157) as a likely rare event in a human sample from a 33 year old male diagnosed with pandemic influenza virus (A/Nagano/RC1/2009(H1N1)) (accession number AB538389) who died from respiratory failure and multiple organ dysfunction syndrome. In the present study, the mutation at position 155 seemed to be common and was not associated with immune status or increase in clinical signs (results not shown). It is possible that this mutation was present in the inoculum virus in low frequency, but that it was replicated more efficiently taking over the most predominant allele identified in the challenge virus. Alternatively, the mutation may have happened in one or several seeder pigs, and then transmitted to the study pigs. This mutation may have provided the virus some advantage when transmitting from the experimentally infected seeder pig to the new hosts but this was outside the scope of this study. Furthermore, none of the other mutations identified were common to both groups.

It is important to highlight in this study that despite the reduced sample size, changes in antigenic sites were not restricted to pigs with immunity indicating that antigenic drift in pigs can happen early during the infection-transmission chain and may be driven independently from immune selection. This must be further investigated given the risk of swine IAV to human health and the impact of IAV on pigs health. Although most of the amino acid changes were observed in antigenic sites of the HA protein considered to be under higher selection pressure, there were changes also observed in the two other regions of HA that stabilize the conformation of the HA protein, the signal peptide and HA2. The biological implications of these substitutions are not known and further studies are needed to characterize their role in pathogenicity and viral fitness. However this suggests that amino acid changes outside the antigenic sites may also be important for the virus to adapt and transmit to new hosts.

In summary this study shows that antigenic drift of IAV can happen in young pigs shortly after infection. Overall, pigs with passive immunity had more nucleotide substitutions compared to naïve pigs. However, in this study these changes did not always result in phenotypic changes in the HA protein that resulted in new antigenic variants. Nevertheless, this study is important in order to highlight the role of the suckling pig as a potential source of IAV genetic diversity in pigs. This is especially important given that extensive movement of suckling pigs takes place after weaning in the swine industry. In addition, nucleotide substitutions that induce amino acid changes were detected in naïve animals as well as in pigs with passive immunity. Some of these changes were able to persist throughout the infection period. Furthermore, changes were also identified in non-antigenic sites indicating that viral adaptation during transmission in pigs is not only dependent on its antigenic characteristics. Overall this study indicates the complexity of genetic diversity in pigs and further studies are needed to understand viral evolution and epidemiology of IAV in swine populations and the risk they represent to people.

Table 8. Tables indicate the pigs in the PASSIV-VAC and NAIVE groups.

Numbers in the top indicate the days post contact (DPC) for each group. Grey boxes indicate the samples that tested positive by RRT-PCR for IAV, and the numbers within the gray boxes indicate the CT value of the samples that were used for HA sequencing. The days where all pigs tested negative were excluded for simplicity (R: replicate, S: seeder).

A

PASSIV-VAC

	ID	0	1	2	3	4	5	6	7	8	9	10
R1	03					28						
	04					28						
	06					29						
	08					27						
	12				27							
	13					27						
	14				26							
	15				26							
	23					26						
	25					26						
	S1	24										
R2	32						30					
	33						29					
	35						32					
	36								31			
	38						29					
	44							30				
	45						32					
	47							29				
	48						33					
	49						29					
	S2	26										
R3	51						27					
	56					29						
	58						30					
	60						28					
	61					31						
	62						32					
	63							31				
	64					30						
	65							29				
	69								29			
	S3	27										

B

NAIVE

	ID	0	1	2	3	4	5	6	7
R1	01					27			
	05					28			
	06					26			
	07			24					
	08				26				
	23					28			
	24				26				
	25						30		
	26				28				
	27				26				
	S4	28							
R2	40						27		
	43						29		
	44					29			
	46							29	
	48					28			
	49							27	
	76					31			
	77						30		
	78						27		
	89				28				
	S5	27							
R3	68						29		
	70				26				
	71				26				
	72				27				
	74					29			
	91				30				
	92				31				
	95						31		
	96				28				
	98						29		
	S6	28							
Other NAIVE infected	S7	29							
	S8	29							
	S9	30							

Table 9. Reciprocal geometric mean HI titers against A/Swine/IA/00239/04 H1N1 (challenge virus) and A/Swine/IL/02450/08 (vaccine strain).

	A/Swine/IA/00239/04 H1N1		A/Swine/IL/02450/08	
Group	Pre-contact	Post-contact	Pre-contact	Post-contact
PASSIV-VAC	17	70	143	111
NAIVE	10	393	21	124

Table 10. Hemagglutinin nucleotide and amino acid identity between A/Swine/IA/00239/04 (challenge virus) and A/Swine/IL/02450/08 (vaccine strain).

	Sequence length	% Identity
HA (nt)	1701	87.9
HA (a.a)	566	88.5
HA1 (nt)	980	87.8
HA1 (a.a)	327	87.5
Antigenic site A*	24	79.2
Antigenic site B	22	72.7
Antigenic site C	33	75.8
Antigenic site D	48	87.5
Antigenic site E	34	82.4
No antigenic sites	166	94.0

nt: Nucleotide sequence

a.a: Amino acid sequence

* Note that antigenic sites are not linear. Each position for each antigenic site previously described was compared between sequences used in this study.

Table 11. Summary of nucleotide substitutions found in the full length HA sequences by group.

Group	Type of substitution[2]	Effect[3]	Nucleotide			Region[7]	Frequency[8]
			Ref.[4]	Pos[5]	Var[6]		
NAIVE (n=16)[1]	Transition	SM	C	282	T	HA1	1
	Transition	NSM	A	15	G	Sig_pep	1
	Transition	NSM	G	515	A	HA1	5
	Transversion	NSM	A	617	C	HA1	1
PASSIV-VAC (n=9)	Transition	SM	T	1194	C	HA2	1
	Transition	NSM	G	515	A	HA1	6
	Transition	NSM	A	563	G	HA1	1
	Transversion	NSM	G	454	C	HA1	1
	Transversion	NSM	A	1382	C	HA2	1

[1]Number of full lengths HA sequences

[2] Type of substitution compared to the reference sequence.

[3] Base pair substitution effect on the hypothetical translated protein (SM: synonymous mutation, NSM: non- synonymous mutation)

[4] Reference nucleotide according to reference sequence

[5] Position based HA numbering

[6] Variant nucleotide

[7] HA region where the substitution was found

[8] Frequency: number of times the specific difference was observed

Table 12. Nucleotide reads in partial sequences at positions where polymorphisms in the HA full length sequences were identified.

Group	Ref[2]	Pos[3]	Read[4]	Freq.[5]
NAIVE (n=12)[1]	C	282	C T N*	10 1 1
	A	15	A	12
	G	515	G A N	7 3 2
	A	617	A N	11 1
PASSIV-VAC (n=16)	T	1194	T	16
	G	515	G A -	11 4 1
	A	563	A -	14 2
	G	454	G -	15 1
	A	1382	A C N	12 1 3

[1]Number of partial HA sequences obtained

[2] Reference nucleotide according to reference sequence

[3] Position based on the full length of the HA sequence

[4] Nucleotide read at the indicated position

[5] Frequency: number of times the specified read was found

Table 13. Non-synonymous (NSM) versus synonymous (SM) mutations in HA full lengths sequences by group.

	Non-Synonymous (NSM)	Synonymous (SM)	Total
PASSIV-VAC	9	1	10
NAIVE	7	1	8
Total	16	2	18

Odds of NSM in the PASSIV-VAC = 9/1 = 9

Odds of NSM in the NAIVE= 7/1 = 7

OR = 1.29 (Fisher's exact test p>0.05)

Table 14. Nucleotide read at position 515 by group.

	Variant = A	Reference = G	Total
PASSIV-VAC	10	14	24
NAIVE	8	18	26
Total	18	32	50

Odds of NSM at position 515 in the PASSIV-VAC = 10/14 = 0.714

Odds of NSM at position 515 in the NAIVE= 8/18 = 0.444

OR= 1.61 (p=0.38)

Table 15. Summary of hypothetical amino acid substitutions found in HA by group.

Group	Region	Amino acid position[1]	Amino Acid		Site[5]	Frequency[6]	Chemical nature
			Ref.[3]	Var.[4]			
NAIVE	H1	NA[2]	NA	NA	-	1	NA
	Sig_pep	5	Isoleucine	Methionine	-	1	Hydrophobic to hydrophobic
	H1	155	Glycine	Glutamate	B	5	Polar to charged
	H1	189	Glutamine	Proline	B	1	Polar to hydrophobic
PASSIV-VAC	H2	NA	NA	NA	-	1	NA
	H1	155	Glycine	Glutamate	B	6	Polar to charged
	H1	171	Lysine	Arginine	D	1	Charged to charged
	H1	135	Alanine	Proline	A	1	Hydrophobic to hydrophobic
	H2	204	Asparagine	Threonine	-	1	Polar to polar

[1]Numbering for each hypothetical protein

[2]NA: No applicable (Synonymous mutation)

[3]Ref: Amino acid predicted in the reference sequence

[4]Var: Amino acid predicted in the variant sequence

[5]Site: Antigenic site where the substitution was mapped

[6]Frequency: number of times the specific difference was observed

[7]Chemical nature of the amino acid substitution

Figure 10. Network analysis of the different alleles identified among all samples.

Each circle represents a different allele and its size is proportional to the number of sequences by allele identified (Total number of sequences included in the analysis = 25). Yellow and blue indicate the sequences from the NAIVE and PASSIV-VAC groups respectively. The numbers indicate the position were nucleotide differences were identified.

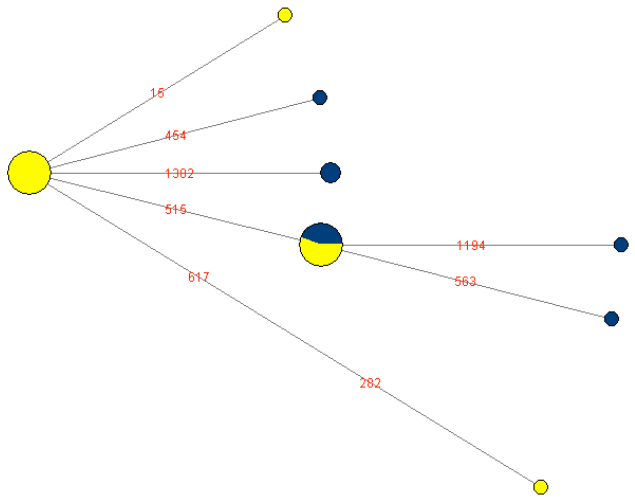
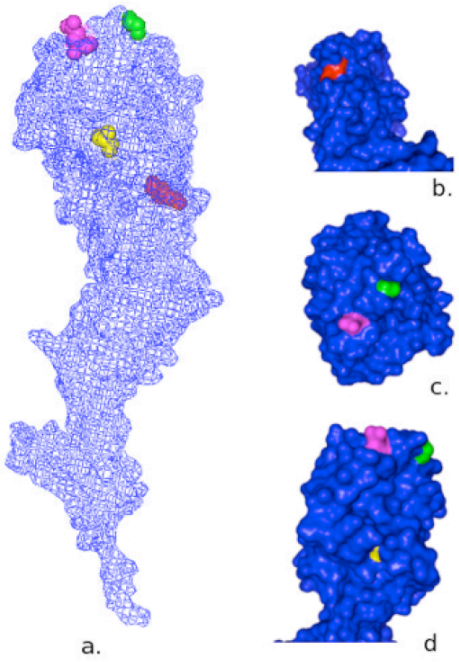


Figure 11. Protein model for one HA1 monomer of A/Swine/IA/00239/04 (challenge virus).

(a) Mesh model highlighting the spatial distribution within the protein of the amino acid that were found different. Close ups b, c and d, represent the predicted surface of the protein and the area exposed for each amino acid replaced. NAIVE: Q189P (magenta) within antigenic site B; PASSIV-VAC: K171R (red) within antigenic site D and A135P (yellow) within antigenic site A. The green spheres represent G155E that was seen in both groups NAIVE and PASSIV-VAC within antigenic site D.



Chapter 5: Genome plasticity of triple reassortant H1N1 influenza A virus during infection of vaccinated pigs

This work has been published in:

Diaz A, Enomoto S, Romagosa A, Sreevatsan S,

Nelson M, Culhane M, Torremorell M. J Gen Virol. 2015. InPress.

Introduction

Influenza A viruses (IAVs) are distributed globally and can infect a wide range of host species including humans (158), birds (2), pigs (16), horses (126), dogs (5), cats (6) and seals (159). Wild waterfowl are considered the natural IAV reservoir (9) and a genetically distinct lineage of viruses has also been identified in bats (8). A swine origin H1N1 IAV was responsible for the first pandemic of the 21st century (10) and was associated with over 200,000 human deaths (160). In recent decades, the genetic diversity of swine IAVs in North America has increased significantly due to the emergence of triple reassortant H3N2 viruses in the late 1990s (54), the numerous introductions of human-origin viruses in pigs including the 2009 pandemic virus (15), and the large-scale movement of pigs between different US regions (13). It is estimated that over 90% of swine herds in the Midwestern US are infected with IAVs (17) and that pigs can be exposed to different IAVs during their life time (23, 128).

IAVs belongs to the family *Orthomyxoviridae* and have a segmented genome composed of eight single-stranded negative-sense RNA segments that encode for at least 12 proteins: polymerase basic 2 (PB2), polymerase basic 1 (PB1), polymerase acid (PA), hemagglutinin (HA), nucleoprotein (NP), neuraminidase (NA), matrix (M) and non-structural protein (NS). RNA viruses have a high mutation rate that increases their genetic diversity over time (161-163), and the segmented nature of the IAV genome allows the virus to exchange (reassort) gene segments with other IAVs contributing to the overall genetic diversity of IAVs.

The main antigenic proteins of the virus, HA and NA, determine IAV subtype. In pigs H1N1, H1N2 and H3N2 are the most prevalent IAV subtypes (3). In North American swine there are six antigenically and phylogenetically distinct H1 groups (α, β, γ1, γ2, δ1, and δ2) (56, 76), and four H3 groups (I, II, III and IV) (54). Multiple IAV subtypes can co-circulate in swine herds and persist at the

population level for prolonged periods of time (17, 128). Additionally, multiple alleles (sequence variants of the same virus) can co-exist during IAV infection of pigs (73) and the same virus can evolve differently in the upper and lower respiratory tract of pigs (74). Furthermore, nucleotide substitutions within the HA antigenic sites can occur shortly after infection of pigs with no significant differences noticed between pigs with or without immunity to the virus (73, 140).

In the US most of the pigs may be exposed to one or more IAVs during their lifetime. Hence the majority of pigs may have immunity to one or several IAVs strains when they are challenged with currently circulating IAVs. However there is a lack of knowledge on how the virus evolves in swine populations that are seropositive to different IAVs. Therefore, the objective of this study was to explore the genetic diversity of the complete genome of a triple reassortant H1N1 IAV population during experimental infection of vaccinated pigs. A vaccine with multiple IAV strains was used to mimic field conditions where pigs are usually exposed to different IAVs. We identified several polymorphisms using next generation sequencing (NGS) directly from nasal swabs and reconstructed 13 complete genomes of the within-host viral populations (metagenomes) using polymorphism overlapping sequence fragments analysis demonstrating the complex, rapid, and dynamic evolution of IAV during infection of vaccinated pigs.

Results

Animal IAV infection status before and after contact with the seeder pig: To study the genomic plasticity of a triple reassortant H1N1 IAV during infection of vaccinated pigs we obtained 11 pigs free of IAV, vaccinated 10 and infected 1 to serve as seeder. All pigs (n=11) were IAV negative by real time RT-PCR (RRT-PCR) and seronegative by NP-ELISA prior to vaccination. Two weeks after the booster vaccination, and before contact with the seeder pig, nine pigs tested ELISA positive to IAV (S/N< 0.6) and 1 was a suspect (S/N=0.803). Three pigs

were negative (HI titer < 1:20) and 7 had HI titers ≤ 1:40 to the challenge virus. All vaccinated pigs had HI titers ≥ 1:20 to the vaccine viruses (Table 16).

The seeder pig remained negative before challenge and tested IAV positive by RRT-PCR 48 hours after challenge. After introduction of the seeder pig into the isolation unit with the remaining pigs (n=10), 5 animals tested IAV positive and 5 tested negative by RRT-PCR during the study period (Table 17). Prior to the introduction of the seeder pig, the mean S/N ELISA titer was not statistically different ($p=0.75$) between pigs that tested positive and negative by RT-PCR. However, 14 days after the introduction of the seeder pig the S/N ELISA titer was lower ($p=0.04$) in pigs that tested RT-PCR positive compared to pigs that tested RT-PCR negative (Table 16). Additionally, we also found statistically significant differences ($p<0.05$) in the HI titers before and after contact to the seeder pig, between pigs that tested positive and pigs that tested negative by RT-PCR (Table 16).

Extensive allelic variation was identified by sequence analysis during infection: The complete genome of IAV was amplified and sequenced from 13 samples which included the inoculum virus before challenge, two samples from the seeder pig at days 2 and 4 (SD2 and SD4) and 2 samples from each of the 5 infected pigs after contact (A1D5, A1D7, A2D4, A2D8, A3D5, A3D6, A4D3, A4D6, A5D4, and A5D6; where, 'A' refers to animal number and 'D' refers to day of the study). The most frequent allele in the inoculum virus for segments 1, 4, 5, 7, and 8 was also the most frequent allele in the pig samples analyzed. In contrast, the most frequent allele for segments 2, 3, and 6 was different in pig samples compared to the inoculum virus. Many nucleotide polymorphisms (n=794) were found in all samples throughout the course of the study and were distributed in all gene segments. However, there was great variability in the number of polymorphisms between samples and gene segments (Table 18). The overlapping sequence fragments analysis estimated a total of 327 alleles of which 214 were unique sequences (Table 19).

Three of the 41 original polymorphisms present in the inoculum virus were not identified in any of the pig samples analyzed and not all alleles in the inoculum were identified in the pig samples sequenced. Moreover only 4 emergent alleles (defined as alleles not present in the inoculum virus) were found in multiple pigs (2 in segment 2, and 1 in segment 3 and 4). Finally, for all but the HA segment, the crude ratio of synonymous to non-synonymous substitutions (ds/dn) was greater than 1 (Table 19).

Although there was small or no variation in the number of alleles detected between the two samples sequenced of most pigs that tested positive, the allele frequency changed significantly within animal 1 and 4 (Table 19). We noticed that prior to contact to the seeder pig, these two pigs (animal 1 and 4) were negative by HI to the challenge virus, and had the lowest HI titer (1:80) to A/Swine/Iowa/110600/2000(H1N1), which is the vaccine virus closest to the challenge virus at the nucleotide level (Table 16).

HA and NA antigenic proteins diverge independently during infection:
While forty-four different alleles were found in HA, only three alleles were found in NA during this study (Fig 12. The 44 HA alleles identified (Fig 12a), yield 43 different predicted peptides (Table 20). The starting virus inoculum contained two HA alleles that only differed in one nucleotide within the HA2 region. In the seeder pig on day 2 (SD2), we found 8 HA alleles (Fig 12a) with variations in the amino acid sequence within both the HA1 and HA2 regions (Table 20). However, the emergent alleles of the SD2 were not found in any other pig samples analyzed. All the other HA emergent alleles were unique to an animal except for one that was detected in two pigs (A1D5, A2D8); however, the latter emergent variant (A1D5/A2D8, highlighted in Fig 12a and Table 20) contained a HA1 region identical to the inoculum alleles. Additionally, we found 32 HA alleles in A1D7; at the nucleotide level, half of these 32 alleles were closer to an allele identified in A4D6 while the other half was closer to an inoculum allele (Fig 12a). At the protein level, these 32 variants contained polymorphisms in all three

regions of the HA, signal peptide, HA1, and HA2. Overall, amino acid substitutions within HA1 were only found in alleles obtained from SD2 and A1D7 (Table 20); 4 of these substitutions happened within antigen sites previously described for HA subtype H1 (Table 20) and their location and nature are illustrated in Fig 13.

In contrast, at the NA level only three alleles were found during this study and all of them translated the same NA protein. Two of these alleles were present in the starting inoculum virus. The majority inoculum allele was not detected in most samples sequenced (except in A1D7), and the minority allele became fixed in most of the pig samples except for A4D6, in which a third emergent NA allele was found (Fig 12b).

Discussion

To better understand the evolution of IAVs during infection of vaccinated pigs we used deep genome sequencing to compare the viral genetic diversity at two separate sampling points during infection. We demonstrated that genetic makeup of the virus changed in all gene segments as the virus replicated within the group of animals, yielding a complex collection of viral genomes with similar and distinct variants. The infection produced a population of heterogeneous alleles by gene segment (usually ≥ 2) that was dynamic over time. Therefore our results indicate that the genetic heterogeneity of IAVs during infection of partially immune pigs is significant and it might have been underestimated. Under this scenario controlling the transmission of IAVs in pigs is challenging because in natural conditions a large proportion of pigs have maternal or active immunity to different IAVs strains (3, 17), IAVs are endemic in swine populations (16), multiple genetic lineages of the virus can co-circulate in pigs (23, 128), and pigs are moved and mixed in large batches of animals during their production stage (25, 27).

In our study the genetic diversity of IAVs changed dynamically throughout the course of infection. Two samples that corresponded to two different pigs (A1D7 and A4D3) had a higher number of alleles compared to the rest of samples. Since we did not sequence all samples from all pigs, and our sample size was limited, we cannot be certain that a high number of alleles were not present in all pigs at some point during infection. However, our results proved that the diversity of IAVs could change within a vaccinated pig throughout the course of infection. Interestingly, these two pigs (animal 1 and animal 4), were negative by HI to the challenge virus (titer < 1:20) before exposure to the seeder pig, and had the lowest HI against A/Swine/Iowa/110600/2000(H1N1) which is the closest vaccine virus compared to the challenge virus. These findings suggest that in pigs the level of antibodies against IAVs might influence the overall diversity of the virus during infection. However, this observation needs to be corroborated in future studies in particular, in the context of heterologous vaccination to infection. Most vaccines are heterologous to circulating viruses and only provide partial protection to infection; therefore the variability in the immune response to IAV vaccination may influence virus evolution. Individual host factors such as response to social stress, (164), individual host genetics, and animal behavior may also affect the immune response to viral infections. In addition, it remains unclear to what extent different alleles are selected for or whether rapid changes in the viral population are primarily stochastic.

In the samples sequenced in this study, the HA segment was more likely to undergo non-synonymous mutations compared to the remaining segments, including NA. Only one HA emergent allele was found in more than one pig and this allele was identical in the HA1 region to the inoculum virus. In agreement with our results, other studies in pigs have shown that nucleotide substitutions can happen in the HA segment very early after infection in pigs with immunity to IAV, and that allele fixation can happen among infected animals (73, 140). However, the time required for these substitutions to become fixed at a population level is still unknown. Our results are also consistent with a previous

study of the HA1 region of the hemagglutinin indicating that IAVs in pigs are not being transmitted as a single genotype but rather as a population of viruses that may be closely related to each other (73). However, our results also showed that nucleotide substitutions can happen in the signal peptide and the HA2 region of the hemagglutinin, which were not evaluated by Murcia et al. 2012. In addition, other factors in our study such as the group housing conditions, which facilitated greater interaction between pigs, compared to previous studies where pairs of individuals were used to measure intra-host diversity of IAV (63, 73, 126), may have had an effect on the increased overall genetic diversity.

In contrast, we did not find non-synonymous mutations in NA and there was therefore no evidence of coevolution of the HA and NA or epistatic interactions. Although the dominant NA allele in the inoculum was not observed among the majority of samples sequenced, including the seeder, the allele was observed in sample A1D7, indicating that it had likely persisted at low levels during transmission. This genotype “recovery” has been described for other RNA viruses during replication such as polioviruses (165). It is possible that unique alleles found in pigs were present in the inoculum at a low prevalence, and that we were not able to identify them in the inoculum itself. Additionally, unique alleles could have been present in the pig samples that were not sequenced.

Multiple studies have evaluated the intra-host diversity of RNA viruses over time (166-168). In IAVs, this research has focused on HA (4, 5, 126). To our knowledge, our study is the first approach to study the intra-host diversity of the complete genome of IAV during infection of pigs using next generation sequencing (NGS) and our results are comparable to a recent study in children (169). We uncovered an additional layer of complexity in the evolution of IAVs during infection of immune pigs by demonstrating that all IAV gene segments replicate as a population of alleles that may or may not be transmitted. Although our methods were not able to capture reassortment events within hosts, the intra-host diversity observed here certainly provides opportunity for novel reassortant

viruses to emerge. In other species, the evaluation of intra-host reassortment of IAV has shown that two viruses that are closely related to each other reassort at different rates depending on their co-infecting dose (170). Genome reassortment should be further investigated since it is a potential source of genetic diversity to swine IAV. It is not clear to what extent the diversity observed in our study is deleterious and not likely to be transmitted onward over longer time periods in pigs. Further understanding of the intra-host dynamics of co-infection and reassortment remains an important outstanding question in IAV evolution.

The dynamic nature of polymorphisms found in our study highlights that IAV genetic diversity ought to be studied directly in the original biological sample (i.e. nasal swab). The study of genetic diversity and evolution of IAV populations through cell culture might be misleading. The cell culture of IAV leads to loss of diversity as cell culture selects for the fastest growing virus in a new environment. As an example, the frequently used MDCK and VERO cell lines have differential preferences for IAV variants, which has led to selective rescue of specific alleles during serial passages (171). As IAV diversity and population dynamics is complex and shaped by many factors including viral fitness, mutation rate, host factors, and stochastic events that may produce bottlenecks, a better estimate of IAV diversity at the population level can be obtained directly from the original sample. However, amplifying the complete genome from IAV isolates, where the viral concentration is exponentially higher than the original sample, might be easier for certain studies.

It is important to mention different external sources of potential variation and bias including sequencing (172), depth of coverage (106), PCR (98, 173) and sampling or intra-assay bias or error. The platform we used, 454, is mature and errors due to sequencing are considered non-issues since we avoid reads with a Phil's Read Editor (Phred) score <20, polymorphisms in homonucleotide runs, not represented in both strands, and not represented in unique independent sequence runs (98, 174). The High Confidence Difference file (HCDiff) that we

used takes these three precautions into account. The 454 has proven to accurately detect human immunodeficiency virus mutants at a prevalence as low as 0.1% (99). Nevertheless, the variability on NGS reads mapped and depth of coverage throughout the complete genome of IAVs remains an issue to better estimate the genetic diversity of the viral populations (106, 175). Additionally, PCR (especially when the polymerase is stalled) generates in vitro recombinants that inflate and distort the estimates of the number and structure of the true alleles (((173), S., Enomoto unpublished data). To avoid PCR errors we used a high fidelity PCR system that uses a blend of DNA polymerases including one isolated from *Pyrococcus furiosus* (Pfu) which has 3' to 5' exonuclease (proofreading) activity, with a tenfold improved error rate compared to TaqDNA polymerase (176, 177). High fidelity polymerases have shown to enhance the PCR and sequencing conditions (178, 179) improving the accuracy to estimate microbial diversity.

In conclusion, the swine influenza A viral population in an experimental setting was complex. Although we recognize our sampling bias to estimate the complete genetic composition of the viral population during transmission, our findings demonstrate that the diversity of IAV can change dynamically during infection of vaccinated pigs. New sequencing technologies and bioinformatics algorithms might provide more precise estimates in future studies. In this study, the polymorphisms were abundant, dynamic and not limited to HA and NA. Some variants were maintained while others were not identified among the samples sequenced. Direct sampling and deep sequencing allowed us to investigate the dynamic plasticity of IAV population during IAV infection in a small group of pigs. We envision that the plasticity of IAV's genome under field conditions is not less complex since different IAV subtypes can coexist and susceptible animals are continuously introduced into infected populations. Our study emphasizes the need to study IAV evolution directly from the infected host using new generation sequencing approaches, which will help design better strategies to control influenza in animals and people. More studies are needed in order to evaluate

whether the changes observed in this study are due to vaccination or whether they are also found in non-immune pigs.

Materials and methods

Study Design:

Eleven 3-week old specific pathogen free (SPF) piglets, were selected from a serologically IAV negative swine herd and moved to the University of Minnesota animal research units. The IAV negative status was confirmed by testing individual nasal swabs with real time RT-PCR (RRT-PCR) targeting the M gene (84, 85) and serum samples by ELISA (Influenza Ab Test Kit, IDEXX Laboratories Inc., Westbrook, ME, USA) for antibodies against the NP (89).

Viral RNA was eluted using 50 µl of each sample into 50 µl elution buffer using MagMax™ virus RNA isolation kit (Ambion®, USA). AgPath-ID™ One-Step RT-PCR reagent kit (Ambion®, Life technologies, USA) was used to detect IAV. PCR mix containing 5 µl RNA, 12.5 µl 2X buffer, 1.0 µl 25X enzyme mix, 1.67 µl detection enhancer, 5 pmol of each primer and 1.5 pmol of probe was run on a LightCycler® 480 system (Hoffmann-La Roche, Switzerland) at 45°C for 10 min, followed by 95°C for 10 min, and 45 cycles at 94°C for 1 sec and 60°C for 30 sec. Fluorescence was recorded at 60°C and a sample was considered positive if the cycle threshold (CT) was lower than 40. This PCR protocol can detect IAVs in samples containing 200 copies or more of the target amplicon, and has a 100% and 95% diagnostic sensitivity and specificity respectively (84).

Ten pigs were vaccinated a day after arrival and two weeks later with 2 ml of a licensed inactivated trivalent IAV vaccine (FluSure ®, Zoetis Animal Health, New Jersey, USA), containing the δ and γ clusters of H1N1 ((A/Swine/NorthCarolina/031/2005(H1N1) and

A/Swine/Iowa/110600/2000(H1N1) respectively)) and one H3N2 (A/Swine/Missouri/069/2005(H3N2)). Two weeks after the second vaccination, nasal swabs and blood samples were collected from all pigs and tested for IAV by RRT-PCR (85) and ELISA respectively. Additionally, blood samples were tested by hemagglutinin inhibition tests (HI) against the challenge and vaccine viruses before and after infection as previously described (90). The mean ELISA and HI titers were compared between vaccinated pigs that tested RT-PCR positive or negative during this study and considered statistically significant if the p value for the non parametric one-way analysis of variance Kruskal Wallis test was lower than 0.05.

One unvaccinated pig was inoculated with IAV in a separate room to serve as a seeder pig to infect the other pigs. Two ml of $1X10^6$ TCID50/ml A/Swine/IA/00239/2004 H1N1 IAV (GenBank accession: EU139832.1), grown in Madin Darby Canine Kidney (MDCK) cells (92) was used to challenge the seeder pig intranasally and intratracheally. The A/Swine/IA/00239/2004 H1N1 clusters within the β H1 swine IAVs (56). This virus was selected because it has been fully characterized, genetically and antigenically (76), and it has been used in several pathogenesis (81) and transmission studies (38, 64, 140). The challenge virus was 91.5 and 73.7 % identical at the nucleotide level to the H1 γ and δ vaccine virus strains respectively. The infection was confirmed 48 hours later by RRT-PCR and the seeder pig was placed in contact with the rest of the pigs. Nasal swabs were collected from all pigs daily for 14 days into 1.8 ml viral transport media (MEM plus, 2% BSA and 1% penicillin-streptomycin) and an aliquot of the transport media was used for RRT-PCR testing. All pigs were euthanized on day 14 and all procedures for this study were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC) protocol number 0908A71965.

Sample selection, genome amplification and sequence identification:

To explore the within host variability of IAV during infection, two IAV positive samples from each pig were conveniently selected for complete genome amplification and sequencing using NGS technologies (Table 17). Samples with the lowest cycle threshold (Ct) value and best genome amplification were targeted for sequencing. IAV genome was amplified using a modified protocol of Zhou *et al.,* 2009 (103). Briefly, the viral RNA was purified from the swabs using QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA, USA, cat: 52904). IAV cDNA was created from viral RNA using primer MBtuni12(M), ACGCGTGATCAGCRAAAGCAGG, and Superscript III First Strand Synthesis SuperMix (Invitrogen™, Grand Island, NY, USA, cat: 18080-400) cDNA was amplified in a PCR (5 cycles of 94°C 15s, 45°C 30s, 68°C 180s and 31 cycles of 94°C 15s, 57°C 30s, 68°C 180s), consisting of PicoMax High Fidelity DNA Polymerase (Agilent, Santa Clara, CA, USA, cat:600422), MBtuni12M and MBtuni13, ACGCGTGATCAGTAGAAACAAGG. PCR products were verified by gel electrophoresis and purified using QIAquick Spin Kit (QIAGEN, Valencia, CA, USA, cat: 28106). Purified cDNA from the virus inoculum and 12 pig samples (Table 17) were submitted to the Genomics Center at the University of Minnesota for library preparation and 454 sequencing (454 GS-FLX ©, Life Sciences, Roche Diagnostics Corporation, Basel, Switzerland) as described in detail by Ramakrishnan *et al*., 2009 (180).

The 454 inoculum reads were assembled with Newbler 2.6 (Roche Diagnostics Corporation, Basel, Switzerland) using a reference template obtained from GenBank (Table 21) and the inoculum consensus sequence was used as the reference genome (RG, Table 22) to assemble the 454 reads from each pig sample. The polymorphisms present in each sample were extracted from the 454HCDiff.txt files created during each assembly in Newbler 2.6. This file includes only highly confident differences which are defined as variants identified in at least 3 unique reads, and present in forward and reverse reads.

Alleles identification and overlapping reading test: Alleles (sequence variants) were defined as complete functional gene segments identified by aligning overlapping sequence fragments. The Newbler output, 454HCDiff.txt, is a file of sequence alignments surrounding all the high confidence polymorphic loci. A RUBY (181) script was written to test the linkage of two adjacent loci by enumerating the occurrence of the four sequence combinations, consensus-consensus, consensus-variant, variant-consensus, and variant-variant. If greater than 80% of the sequences occurred only as two sequence combinations, the two loci were considered linked. Presence or absence of polymorphisms at each locus was encoded as 1 or 0, respectively. The alleles were deduced by linking together the adjacent intervals between the two polymorphic loci and its functionality verified using the NCBI FLu ANnotation tool (FLAN) (131). Additionally, if the distance that separated two polymorphisms was longer than the length of the reads obtained, then those two polymorphisms were considered not linked. For example if two adjacent polymorphic loci were linked and recovered as 00 and 11, the segment contained two alleles rather than four alleles. The raw 454 reads, the alleles sequences obtained and the Ruby scripts for overlapping sequence fragments analysis and allele extraction are available upon request.

Sequence analysis: To illustrate the phylogenetic relationship between sequences, alleles were aligned to the RG using DNA-Alignment and median-joining networks were estimated using Network (153). Each network was annotated with Network Publisher (Fluxus Technology Ltd, Clare, Suffolk, England) and Adobe Illustrator CC (Adobe Systems Incorporated, CA, USA). Additionally, for the first open reading frame (ORF) we estimated the average number of synonymous (d_S) and non-synonymous (d_N) mutations and their ratio (d_S/d_N) among sequences (182, 183) using the Synonymous and Non-synonymous Analysis Program available at www.hiv.lanl.gov (SNAP, Los Alamos National Laboratory, Los Alamos, NM, USA).

Hemagglutinin and neuraminidase protein analysis:

For HA and NA, hypothetical proteins were inferred from nucleotide sequence, aligned using ClustalX (137) and compared. The amino acid differences among HA sequences were mapped to the known H1 antigenic sites (155, 156), modeled using the tools available at swissmodel.expasy.org (154) and illustrated PyMOL Molecular Graphic Systems, Version 1.5.0.4 Schrödinger, LLC, New York, NY, USA. The HA1 IAV template used for our protein model was A/Swine/Iowa/15/30(H1N1) (Protein Data Bank (PDB) ID:1RUY. This template was used because this virus is from swine origin, the HA has been crystallized, and it is available for public use.

Table 16. Influenza A virus (IAV) serology results by ELISA and hemagglutinin inhibition (HI) tests for pigs prior to start the study (before vaccination), after vaccination and after infection.

For simplicity, animals that tested IAV RT-PCR positive after exposure to the seeder pig (n=5) were renamed A1 to A5 and animals that tested negative were renamed A6 to A10. ELISA results are expressed as sample to negative ratio (S/N) and considered positive, suspect or negative when S/N < 0.6, 0.6<S/N<0.9 or S/N>0.9 respectively. HI titers are expressed as the reciprocal dilutions and considered positive at >1:20. Pigs 1 to 5 tested positive to IAV by RT-PCR and pigs 6 to 10 tested negative. The reciprocal mean titter is compared between pigs that tested RT-PCR positive and negative after exposure to the seeder pig and considered statistically significant if the p value for the Kruskal Wallis test was lower than 0.05

ID	PCR Result	Before vaccination	Two weeks after second vaccination (Prior contact with the seeder pig)					Two weeks after infection				
		ELISA	ELISA	Hemagglutinin inhibition test*				ELISA	Hemagglutinin inhibition test			
				00239	31	110600	MO069		239	31	110600	MO069
Seeder	Positive	0.967	0.869	Neg	Neg	Neg	Neg	0.32	80	160	0	40
A1	Positive	0.963	0.503	Neg	20	80	160	0.194	320	320	80	160
A2	Positive	1.700	0.803	20	20	160	320	0.129	640	640	80	320
A3	Positive	0.979	0.331	Neg	40	160	320	0.135	160	160	80	320
A4	Positive	0.925	0.332	Neg	20	80	320	0.202	160	320	80	320
A5	Positive	1.009	0.362	20	40	320	640	0.164	640	640	320	640
Mean		1.115	0.466	10	28	160	352	0.165	384	416	128	352
A6	Negative	0.958	0.546	40	40	320	640	0.434	20	80	320	640
A7	Negative	0.968	0.187	20	40	320	640	0.219	40	80	320	640
A8	Negative	1.018	0.545	20	20	160	320	0.247	20	40	80	320
A9	Negative	0.968	0.430	40	20	640	640	0.275	40	80	160	640
A10	Negative	0.981	0.193	20	40	320	640	0.162	20	80	80	320
Mean		0.979	0.380	28	32	352	576	0.2674	28	72	192	512
p value Krusal Wallis test		0.91	0.75	0.03	0.55	0.048	0.06	0.04	0.03	0.007	0.28	0.16

*IAV isolates used in the HI test:

00239: Challenge virus, A/Swine/IA/00239/2004 (H1N1)

031: Vaccine virus: A/Swine/NorthCarolina/031/2005(H1N1)

110600: Vaccine virus: A/Swine/Iowa/110600/2000(H1N1)

M0069: Vaccine virus: A/Swine/Missouri/069/2005(H3N2).

Table 17. Real time influenza A virus RT-PCR results and samples selected for deep genome sequencing.

Nasal swabs were collected and tested for 14 DPC (days post contact). No pig tested positive after 8 DPC. The seeder pig is identified as S. Results from five pigs that did not test positive at any point during this study are not shown. The days when pigs tested positive are indicated by "+" (plus sign) and the RRT-PCR cycle threshold (C_t) value for each sample is shown. A total of 12 pig samples were selected for complete genome sequencing: Seeder pig, day 2 (SD2, dark blue) and day 4 (SD4, light blue); animal A1, day 5 (A1D5, dark green) and day 7 (A1D7, light green), animal A2, day 4 (A2D4, purple) and day 8 (A2D8, light purple); animal A3, day 5 (A3D5, brown) and day 6 (A3D6, light brown); animal A4, day 3 (A4D3, dark grey), and day 6 (A4D6, light grey); animal A5, day 4 (A5D4, pink), and day 6 (A5D6, light pink). Sequences obtained from each sample are color coded in Fig 12 according to the colors indicated in this table.

	Days post contact									
Animal	1	2	3	4	5	6	7	8	9	10
S	+ 27.2	+ 26.0	+ 27.9	+ 33.4						
A1				+ 30.3	+ 27.3	+ 30.2	+ 33.6			
A2			+ 31.0	+ 33.1	+ 29.1	+ 31.2		+ 31.8		
A3				+ 27.6	+ 27.8	+ 29.9	+ 29.5	+ 34.3		
A4			+ 32.9	+ 30.8	+ 27.6	+ 29.2	+ 33.4			
A5			+ 30.6	+ 30.7	+ 28.2	+ 29.1	+ 34.2			

Table 18. 454 reads assembly statistics by sample indicating the total number of reads mapped, depth of coverage and polymorphisms found among samples.

	Inoculum	SD2	SD4	A1D5	A1D7	A2D4	A2D8	A3D5	A3D6	A4D3	A4D6	A5D4	A5D6	Average
Total reads	61363	14277	84236	20623	43769	49677	28386	27365	54858	24760	32666	40468	68539	42383.6
Total mapped	3339	8070	80289	16415	12391	46513	25475	21582	40557	17206	20995	32912	45272	28539.7
Ave. depth	102.9	229.4	2527.3	493.8	291.6	1454.6	718	689.5	1244.4	529.1	641.5	1023	1384.6	871.515
Ave. map length	489	374	417	397	318	413	375	420	403	407	407	408	400	402.154
All differences	168	323	588	308	1184	519	454	397	421	529	280	521	661	488.692
HCDiff.	41	35	12	14	506	13	16	12	13	94	15	12	11	61.0769
HCDiff. %	24.40	10.84	2.04	4.55	42.74	2.50	3.52	3.02	3.09	17.77	5.36	2.30	1.66	12.50
Ave. depth variant	17.4	43.1	303.2	149.35	58.2	232.4	154.4	203.5	215.5	31.9	137.3	258.2	316.5	163.15

HCDiff: Highly confident differences (considered true polymorphisms). Total among all samples sequences: 794

HCDiff. %: Percentage of HCDiff found among all differences.

There was no linear association ($p>0.05$) between the following comparisons:

1. Total mapped and all differences
2. Total mapped and highly confident differences
3. Ave. depth and all differences
4. Ave. depth and all highly confident differences

There was a negative linear association (-0.18, $p=0.001$) between average map length and highly confident differences. However, the average map length only explained 41% of the variability in the confident differences ($R^2=0.415$).

Table 19. Number of alleles distributed by sample and gene segment.

The last column of the table indicates the ratio between synonymous mutations (ds) and non-synonymous mutations (dn) for each gene segment. 214 out of 327 alleles found were unique sequences.

Segment	Inoculum	SD2	SD4	A1D5	A1D7	A2D4	A2D8	A3D5	A3D6	A4D3	A4D6	A5D4	A5D6	Unique sequences	ds/dn
1 (PB2)	4	1	2	2	0	1	2	1	1	4	2	1	1	11	15.7
2 (PB1)	2	2	1	2	1	2	2	2	2	1	1	1	1	5	3.7
3 (PA)	8	2	2	4	0	2	1	2	0	64	2	2	2	77	12.5
4 (HA)	2	8	2	2	32	2	2	1	1	1	2	2	1	44	0.8
5 (NP)	16	2	2	2	8	2	2	2	8	4	2	2	2	31	10.9
6 (NA)	2	1	1	1	2	1	1	1	1	1	2	1	1	3	N/A
7 (M)	2	2	1	1	4	1	1	1	1	1	1	1	1	7	41.0
8 (NS)	2	1	1	1	32	1	1	1	1	4	1	1	1	36	6.1
Total	38	19	12	15	79	12	12	11	15	80	13	11	10	214	

N/A: No dn mutations were found in segment 6, hence no ds/dn ratio was estimated

Table 20. ClustalX alignment of the complete hypothetical HA proteins found by sample.

Only polymorphic sites among alleles identified in this study are shown. Superscripts A, B and D indicate the antigenic site where changes in HA1 were observed. The first two rows indicate amino acids found in the vaccine viruses at the polymorphic sites identified in the samples sequenced. The reference amino acid (RG) for each polymorphic position is shown in inoculum allele 1. Non-highlighted proteins are unique variants and proteins highlighted with the same color are 100% identical among them. A total of 58 functional HA sequences were identified among all samples. These sequences represented 44 unique alleles and translated 43 different hypothetical HAs.

		Protein region and amino acid position											
		Sig_pep	HA1						HA2				
Sample ID	Allele	17	25	29	159[A]	180[D]	200[B]	252[D]	369	370	371	372	373
IAV031	NA	A	H	S	S	F	N	P	NA	NA	NA	NA	NA
IAV110600	NA	A	H	S	N	K	S	E	H	H	Q	N	E
Inoculum	1	A	H	S	N	K	P	E	H	H	Q	N	E
	2	.	.	.	.	.	.	.	.	.	.	-	K
SD2	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	.	.	.	-	K
	3	.	.	.	.	.	.	K	.	.	.	.	.
	4	.	.	.	.	.	.	K	.	.	.	-	K
	5	.	.	.	.	.	L	.	.	.	.	.	.
	6	.	.	.	.	.	L	.	.	.	.	-	K
	7	.	.	.	.	.	L	K	.	.	.	.	.
	8	.	.	.	.	.	L	K	.	.	.	-	K
SD4	1*	.	.	.	.	.	.	.	.	.	.	.	.
	2*	.	.	.	.	.	.	.	.	.	.	.	.
A1D5	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	.	.	.	Q	K
A1D7	1	.	R	P	.	.	.	.	.	.	.	.	.
	2	.	R	P	.	.	.	.	.	.	-	Q	K
	3	.	R	P	.	.	.	.	Q	Q	-	.	.
	4	.	R	P	.	.	.	.	Q	Q	-	-	K
	5	.	R	P	K	.	.	.	.	.	.	.	.
	6	.	R	P	K	.	.	.	.	.	.	-	K
	7	.	R	P	K	.	.	.	Q	Q	-	-	K
	8	.	R	P	.	E	.	.	Q	Q	-	.	.
	9	.	R	P	K	E	.	.	.	.	.	.	.
	10	.	R	P	.	E	.	.	.	.	.	.	.
	11	.	R	P	K	.	.	.	Q	Q	-	.	.
	12	.	R	P	K	E	.	.	Q	Q	-	.	.
	13	.	R	P	.	E	.	.	.	.	-	Q	K
	14	.	R	P	K	E	.	.	.	.	-	Q	K
	15	.	R	P	K	E	.	.	Q	-	-	Q	K
	16	.	R	P	.	E	.	.	Q	-	-	Q	K
	17	G	R	P	.	E	.	.	.	.	-	Q	K
	18	G	R	P	.	E	.	.	Q	-	-	Q	K
	19	G	R	P	K	E	.	.	Q	-	-	Q	K
	20	G	R	P	K	E	.	.	.	.	-	Q	K
	21	G	R	P	K	E	.	.	.	.	.	.	.
	22	G	R	P	.	E	.	.	.	.	.	.	.
	23	G	R	P	.	E	.	.	Q	Q	-	.	.
	24	G	R	P	K	E	.	.	Q	Q	-	.	.
	25	G	R	P	.	.	.	.	Q	Q	-	.	.
	26	G	R	P	K	.	.	.	Q	Q	-	.	.
	27	G	R	P	K	.	.	.	Q	Q	-	-	K
	28	G	R	P	.	.	.	.	Q	Q	-	-	K
	29	G	R	P	K	.	.	.	.	.	.	-	K
	30	G	R	P	K	.	.	.	.	.	.	.	.
	31	G	R	P	.	.	.	.	.	.	.	-	K
	32	G	R	P	.	.	.	.	.	.	.	.	.
A2D4	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	.	Q	.	-	K
A2D8	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	.	.	.	Q	K
A3D5	1	.	.	.	.	.	.	.	.	.	.	.	.
A3D6	1	.	.	.	.	.	.	.	.	.	.	.	.
A4D3	1	.	.	.	.	.	.	.	.	.	.	.	.
A4D6	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	Q	Q	-	-	K
A5D4	1	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	.	.	.	.	.	.	.	.	-	K
A5D6	1	.	.	.	.	.	.	.	.	.	.	.	.

*Indicate identical proteins within a sample translated from two different alleles.

NA: Not applicable. (The complete sequence for A/Swine/NorthCarolina/031/2005(H1N1) is not available)

Table 21. Influenza A virus template sequences used to assemble inoculum reads.

Gene segment	GenBank accession number
1 (PB2)	CY099076.1
2 (PB1)	CY099309.1
3 (PA)	CY045233.1
4 (HA)	FJ789832.1
5 (NP)	CY009919.1
6 (NA)	GU236519.1
7 (M)	DQ150436.1
8 (NS)	CY050162.1

Table 22. A/swine/Iowa/00239/2004(H1N1) reference genome (RG) used to assemble all pig sample reads, and to compare allele sequences found during this study.

Gene segment	GenBank accession number
1 (PB2)	KM198687
2 (PB1)	KM198688
3 (PA)	KM198689
4 (HA)	KM198690
5 (NP)	KM198691
6 (NA)	KM198692
7 (M)	KM198693
8 (NS)	KM198694

Figure 12. Median joining networks of Hemagglutinin (HA) and Neuraminidase (NA) alleles found during experimental IAV infection of vaccinated pigs.

Each circle represents a sequence variant (allele) and each color represents the sample where that sequence was found. Red numbers indicate the number of nucleotide differences between sequences (not all numbers are included for brevity). Within each network, the branch length is proportional to the number of differences between alleles. (*) Indicates an emergent HA allele (not present in the inoculum virus) that was found in more than one pig. The color code of this Fig coincides with the color code of Table 17

a. Segment 4 (HA)
44 unique alleles

b. Segment 6 (NA)
3 unique alleles

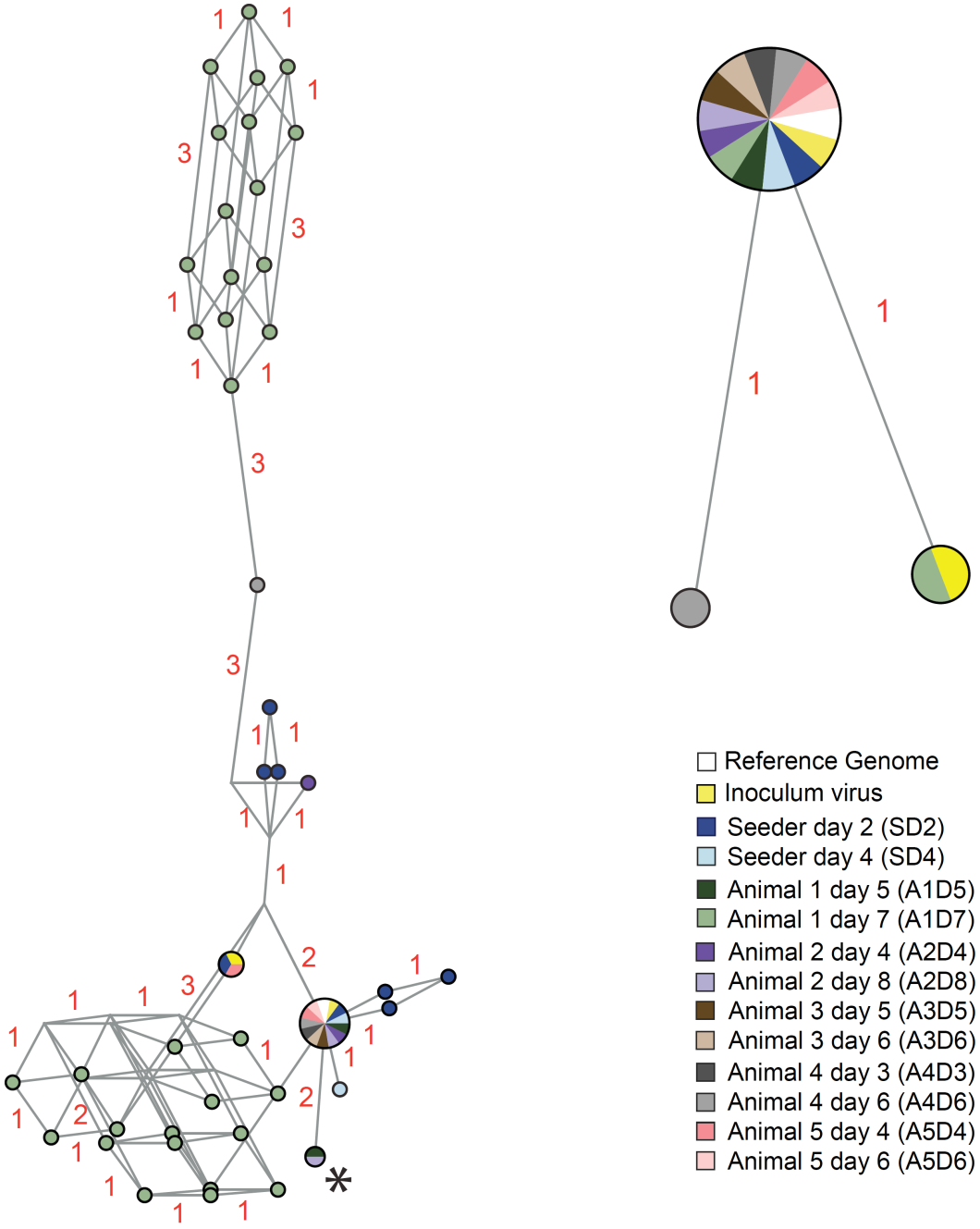
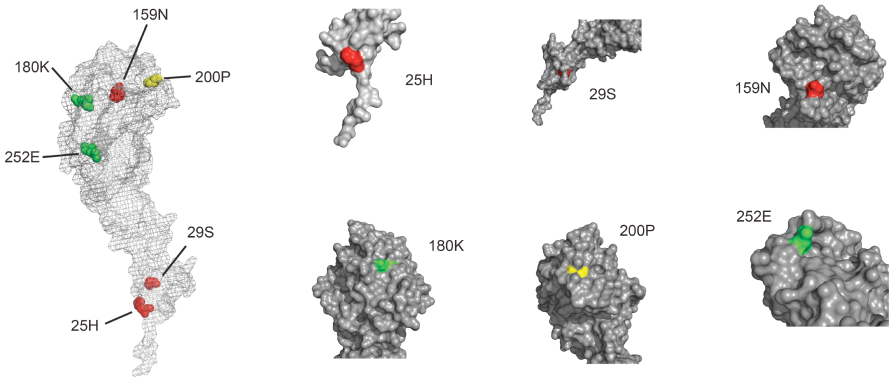
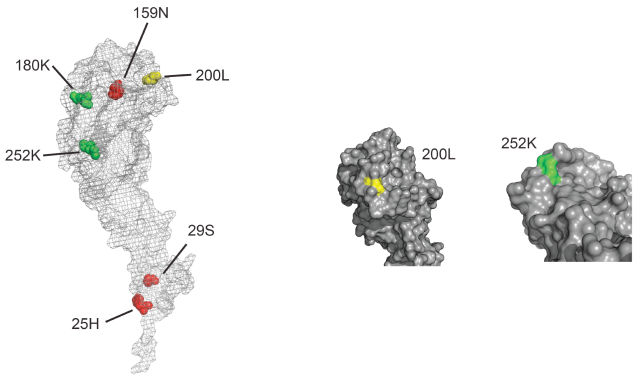


Figure 13. Three-dimensional models illustrating the HA1 region of the hemagglutinin and the polymorphic amino acids found during the study.

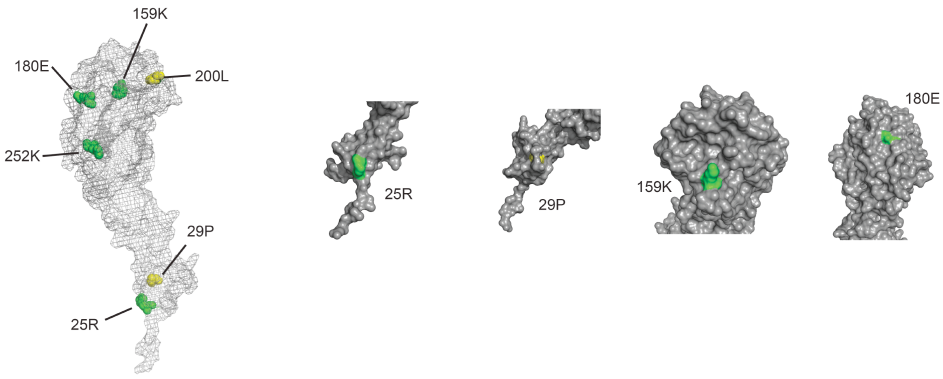
Polymorphic amino acid (histidine (H), serine (S), asparagine (N), lysine (L), proline (P), glutamate (E), leucine (L), arginine (R), and proline (P)) are indicated and colored according to their physical properties: polar (red), charged (green) and hydrophobic (yellow). a) Reference amino acids at polymorphic sites in HA1 for A/Swine/IA/00239 (challenge virus) b) amino acids residues predicted from the sample sequenced from the seeder pig at day 2 (SD2) c) amino acids residues predicted from the sample sequenced from animal 1 at day 7 (A1D7).



a. HA1 A/Swine/IA/00239/2004



b. Seeder day 2



b. Animal 1 day 7

Chapter 6: A prospective cohort study and deep genome sequencing demonstrate the complexity of infection, re-infection and molecular evolution of swine influenza A viruses in pigs

Introduction

Influenza A viruses (IAVs) are Orthomyxoviruses with eight single stranded negative sense RNA gene segments namely polymerase B2 (PB2, segment 1), polymerase B1 (PB1, segment 2), polymerase A (PA, segment 3), hemagglutinin (HA, segment 4), nucleoprotein (NP, segment 5), neuraminidase (NA, segment 6), matrix (M, segment 7), and non-structural protein (NS, segment 8). The main antigenic proteins of IAVs (HA and NA) are used to classify the virus into different subtypes and include at least seventeen HAs (H1-H17(8, 45, 123)) and 9 NAs (N1-N9 (2, 123)). To date at least 116 HA-NA IAV combinations have been isolated from avian species and aquatic birds are considered the natural reservoir for most IAVs present in nature (9). IAVs can also infect different mammalian species including humans (184) and pigs (107) although few IAV subtypes are endemic in these two species. In North America, H1N1, H1N2, and H3N2 IAVs are endemic in pigs and cluster in six H1 (α, β, γ1, γ2, δ1, and δ2) and four H3 (I, II, III and IV) genetic lineages (54, 55, 76).

Multiple reassortment events and the introduction of human IAVs, including the 2009 pandemic virus, changed the genetic landscape of swine IAVs in North America (16, 54, 75). The international trade of live swine and pig movement within the USA are associated with the increased genetic diversity of swine IAVs (13, 24). However, the ecological mechanisms that led to the current genetic diversity of swine IAVs in North America are not clearly understood. There are several reports of distinct IAVs coexisting in pig populations and the same IAV subtype can be found within the same population for prolonged periods of time (17, 23, 37). Moreover, in pigs the viral genome of IAV can replicate during infection as a dynamic “cloud” of genotypes closely related to each other (141) illustrating the complexity of IAV genetic evolution during infection of pigs. Nevertheless, there is a lack of knowledge regarding the molecular epidemiology and genetic diversity of swine IAVs under field conditions, although it is common

to find pigs that are exposed to one or more IAVs during their lifetime (114, 120, 122, 185).

In the USA, the swine industry produces pigs in batches that are typically moved between different production sites before harvest (127). Pigs are born in swine breeding herds and weaned at approximately 3 weeks of age to a nursery or to a wean-to-finish farm where they stay for seven weeks or until market, respectively. If pigs are weaned into a nursery farm then they are moved to a finishing farm where they stay until harvest. Moreover, pig production flows can be managed as continuous or all-in/all-out flows. In a continuous flow, pigs are moved in and out of the production unit (e.g. site, farm, building) on a regular basis and there are pigs all the time within the unit. In an all-in/all-out flow, a batch of pigs is used to fill the unit (room, barn, or site), and no more pigs are added until all pigs from the previous batch have left. In both types of pig flows pens and crates are cleaned between batches. Nevertheless, under these production systems controlling the transmission of IAVs is challenging and requires a better understanding of the epidemiology and molecular biology of swine influenza, which is a main cause of respiratory disease in pigs. Additionally, IAVs containing genes similar to IAVs circulating in North American and Asian swine caused the 2009 IAV pandemic (10) after reassorting in an unknown host; hence, understanding the epidemiology and molecular evolution of swine IAVs is also important to minimize its zoonotic potential.

We hypothesize that there are three possible explanations for the persistence and genetic diversity of IAV in pig herds after weaning. First, there may be differences between IAVs from the same genetic lineage that enable the virus to evade herd immunity over time. Second, weaned pigs may be the source of the same IAV lineage and could introduce similar viral variants with each group of pigs weaned; and third, new weaned pigs may allow the continuous replication of IAVs present in the swine herd. Therefore, the objectives of this study were to assess the epidemiological characteristics and molecular traits of IAVs in a large

group of naturally infected 3-week old pigs for a prolonged period of time. We designed a prospective cohort study with intensive sampling events in pigs after weaning and characterize the complete genome of the viral population over time under natural conditions of transmission. We found that different IAVs produced two overlapping epidemic waves of infection and that the majority of pigs became infected and re-infected during the study period. These results are important because we proved the complex dynamics of IAV evolution and diversity in pigs after weaning and demonstrated the re-infection of pigs with IAVs that are closely related to each other. Understanding the epidemiology and molecular diversity of IAVs in pigs is important to design effective health interventions that aim to reduce the impact of IAVs on swine health and production, and might reduce the public health risk associated with swine IAVs.

Materials and methods

Protocols and procedures followed throughout the study were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC 1207B17281), and the Institutional Biosafety Committee (IBC 1208H18341).

To estimate the patterns of IAV infection in pigs after weaning and characterize the genetic diversity of the virus throughout the growing pig period we designed a prospective, randomized pig cohort study in a wean-to-finish farm. This farm had a single source of weaned pigs and history of IAV infection. Additionally, this farm housed pigs from weaning (3 weeks of age) to market (23 to 24 weeks of age) in eight different barns, with an all-in/all-out pig flow by barn. The cohort of pigs (n=132) was selected from ~2200 pigs that were weaned into a single barn on three different days within a week. Pigs in the cohort were randomly selected from the first shipment of pigs. A random number was assigned to each pig at arrival and the pigs with the lowest random number were selected as the cohort

to study. Sample size was estimated to be 95% confident to detect at least one IAV positive pig if the prevalence of infection was 2.5% or higher. Each pig in the cohort was ear tagged, individually identified and classified as male or female. We kept the cohort of pigs comingled among all other pigs in the barn (as distributed by the pig farmer after arrival) and collected individual nasal swabs (BBL CultureSwab, Becton Dickinson and Company, USA) on a weekly basis for 15 weeks. Every week after sample collection, swabs were refrigerated and transported to the laboratory within 6 hours of collection on the manufacture's transport media and then placed into 1.8 ml sample storing media within 24 hours of collection (Dulbecco's Modified Eagle Medium (DMEM), 5% antibiotic-antimycotic (Gibco, Life Technologies, USA containing 10000 IU/ml of penicillin, 10000 µg/ml of streptomycin, and 25 µg /ml of Fungizone), and 2% bovine serum albumin (BSA) fraction V 7.5% solution (Gibco, Life technologies, USA). Swabs in the sample storing media were vortexed for 10 seconds and stored at -80ºC until IAV testing.

All nasal swabs were tested individually by reverse transcriptase real time polymerase chain reaction (RRT-PCR) for swine IAVs using protocols described elsewhere (84, 85). A sample was considered positive if the RRT-PCR cycle threshold (Ct) value was 35 or lower. The weekly prevalence rate (number of positive cases per week among all pigs tested) and the period prevalence (number of pigs that tested positive at least once during the 15-week study period) were estimated. The number of prevalent cases was compared between weeks and considered statistically significant if the p value for the McNemars test was lower than 0.05. Additionally, the weekly incidence density (number of new cases per pigs at risk during a week) was estimated and the incidence density ratio between males and females compared. The incidence density ratio between males and females was considered significant if the test-based 95 % confidence interval did not include 1. Additionally, we estimated the number of weeks that the same pig tested IAV positive and defined a "re-infection case" as a pig with two or more positive samples in non-consecutive weeks.

To characterize the genetic diversity of IAV during infection of pigs after weaning a set of 92 RT-PCR positive swabs was selected. Sample selection targeted those pigs with 2 or more positive swabs during the study period and those samples with the lowest Ct value. At least one positive sample per week was included. IAV genome was amplified directly from the nasal swab in a single reaction using methods previously described (103). Briefly, viral RNA was extracted from positive swabs using MagMax Viral RNA isolation kit (Ambion, Life Technologies, USA). One step RRT-PCR was performed using SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen, Life Technologies, USA). A 100 µl PCR mix was prepared containing 20 µl DNase/RNase-Free distilled water (Gibco, USA), 50 µl 2x reaction mix, 2 µl SuperScript III RT mix, 2 µl (10 µM) of each primer (MBtuni12(M): ACGCGTGATCAGCRAAAGCAGG and MBtuni13: ACGCGTGATCAGTAGAAACAAGG), and 24 µl of RNA template. Gel electrophoresis was used to verify visually PCR amplicons. PCR products were purified using QIAquick Spin Kit (QIAGEN, USA), and eluted in 20 DNase/RNase-free distilled water (Gibco, Life Technologies, USA). Samples were then submitted to the University of Minnesota Genomics Center (UMGC) for library preparation (TruSeq DNA HT sample prep kit, Illumina, USA) and sequencing using next generation sequencing (NGS) technologies (MiSeq paired end 250 cycles, Illumina, USA).

Sequencing quality control and quality assurance (QC/QA) was first verified with FastQC (129) and then Trimmomatic (102) was used to trim low quality reads using the pair-end mode of the software. Sequencing assembly was performed using Bowtie2 (100) and SAMTools (97) on a reference template containing 6 IAV internal gene references and 4 antigenic gene references (PB2 (CY099076.1), PB1 (CY099309.1), PA (CY045233.1), NP (CY009919.1), M (DQ150436.1), and NS (CY050162.1), H1 (FJ789832.1), H3 (KC992248.1), N1 (GU236519.1), N2 (KC866483.1)). If more than one IAV genotype was found

during the study then different genome templates, obtained from the samples sequenced, were used to re-map the reads of all samples. The proportion of Illumina reads that were mapped to each IAV template was estimated by gene segment and week. Consensus sequences for each gene segment of IAV were trimmed to coding regions and their functionality verified using the NCBI FLu Annotation web-service (FLAN (131)). Complete functional sequences from this final assembly were used to estimate IAV diversity during the study period. Furthermore, complete genomes were called when the eight complete gene segments were assembled. Mixed IAV infections were defined as samples where two or more complete consensus sequences for the same gene segment were obtained after mapping the Illumina reads to multiple IAV genome templates.

The main antigenic gene segments (HA and NA) were analyzed first and the antigenic subtype determined. Then, HA sequences were used to classify IAV into different viral groups (VG) based on the phylogenetic origins of H1 (gamma 1, gamma 2, delta 1, delta 2, alpha, or beta (76)) or H3 (clusters I to IV (54)) subtypes using the web-based tools available at the Influenza Research Database (132). The number of virus variants for each VG was determined and the persistence of the same VG at the population level was estimated over time. Then all other gene segments were classified based on the HA designated VGs and the genome constellation for each sample sequenced was inferred. Additionally, each set of IAV gene sequences (gene segments 1 to 8) was aligned using ClustalW (137) and the pairwise distance identity was used to compare the viral diversity across gene segments using heat maps. Sequencing assembly and analysis was done using the resources available at the University of Minnesota Supercomputing Institute (MSI). Statistical analysis and heat maps were performed using tabular methods and ggplots 2.17.0 in R.

Hypothetical HA and NA proteins were translated and amino acid sequences compared using median-joining network analysis (153). First, all protein sequences were aligned to the most frequent sequence found during the study

period using DNA-Alignment (Fluxus Technology Ltd, Germany). Then, polymorphic sites among antigenic proteins were estimated, polymorphic amino acids per site inferred and median-joining networks constructed (153). Protein networks were drawn and annotated using Network 4.613, Network Publisher (Fluxus Technology Ltd, Germany), and Adobe Illustrator CC 2014 18.1.1 (Adobe Systems Incorporated, USA).

Finally, the patterns of IAV infection and re-infection were compared among those pigs of which we sequenced more than one positive sample during the study period. The percent sequence identity of HA at the nucleotide level was used to estimate virus divergence within pigs over time and the findings compared to the different HA amino acid sequences found during the protein prediction analysis.

Results

A cohort of 132 3-week old pigs (77 males (58.3%) and 55 females (41.7%)) was randomly selected from a group of weaned pigs at weaning (week 0 (W0)) to a commercial wean-to-finish-farm. Males and females were housed on the right and left side of the barn respectively and pigs of different sex did not have nose-to-nose contact. Pigs in the cohort were comingled at arrival with the remaining weaned pigs in the batch and were followed for 15 weeks (W1 to W15). While no selected females died during this study, five males (6.5%) died at weeks 3, 7, 8, 11 and 14; unfortunately, their cause of death was not determined. A total of 2,080 individual nasal swabs were collected over 15 weeks and 369 (17.7%) of them tested positive for IAVs by RRT-PCR. At weaning (W0), 27 pigs (20.5%) tested IAV RRT- PCR positive without a significant difference between males (n=17) and females (n=10) (p=0.74). The 15-week period prevalence of IAVs infection was 98.4% (n=130). Only two pigs (1.6%) tested IAV negative throughout the study period, although one of these pigs died in W3. Moreover, 116 pigs (87.9%) tested positive to IAV more than once (Fig 14) and 103 pigs

(78%) were IAV positive in at least two non-consecutive weeks, hence considered re-infected with IAVs.

There were two epidemic waves of IAV infection during the study period (Fig 15) with no statistical difference between the number of prevalent cases at the epidemic peaks in W2 and W7 ($p=0.24$). Within 15 weeks, the weekly prevalence of IAV infection ranged between 0% and 65.2% (Table 23). The number of IAV positive males and females was statistically different only at W1 ($p=0.04$), W2 ($p=0.004$), and W6 ($p=0.04$). However, given the distribution of pigs within the barn (males on the right and females on the left) it was not possible to determine if this difference was associated with the pigs' sex (castrated male or female) or their location within the barn (right or left). The incidence density of IAV infection of pigs after weaning ranged between 0 and 71 cases per 100 pig-week (Table 23) and for most of the weeks it was not statistically different between males and females. However, at week 6 the incidence density of IAV infection in males was 3 times higher than in females ($p<0.05$).

Ninety-two out of 369 positive swabs (25%) were used for IAV genome amplification and deep genome sequencing using Illumina as the next-generation sequencing (NGS) platform. After the first template-assembly, three different influenza A virus groups (VG) were identified (VG1, VG2, and VG3). At the HA level VG1, VG2 and VG3 clustered within North American H1-gamma, H1-beta and H3-cluster-IV IAVs respectively. At the NA level VG1 and VG2 were N1 viruses while VG3 were N2 viruses. The pairwise percent identity (ClustalW) between the consensus sequences of all gene segments of three representative samples (n=24) of the three VGs is shown in Table 24. These 24 templates were used to re-assemble all samples sequenced. After quality control and quality assurance, nine samples (10.8%) yielded only partial IAV gene contigs and were excluded from the analysis. From the remaining samples (n=83), 13,559,009 Illumina sequencing reads were successfully mapped to the reference templates (Fig 16). Overall, the majority of Illumina reads obtained from the first epidemic

wave of IAV infection mapped to the reference templates of VG1 (H1 gamma virus) while the majority of reads obtained after week 6 mapped to the reference templates of VG3 (H3 cluster IV virus). However, in most weeks we detected reads that mapped to at least two different VGs. Illumina reads mapping to the reference templates of VG2 only predominated in samples sequenced in W4 and W14 (Fig 16). The sequencing assembling process yielded 649 complete IAV sequences, from 83 pig nasal. These 649 sequences were classified based on the sequence template mapped (Table 25) as VG1 (n=402, 61.9%), VG2 (n=19, 2.9%), and VG3 (n=228, 35.1%). There were no complete IAV gene segments obtained from W9 to W13 and W15.

We found different genome constellations among all samples sequenced (Fig 17) and different VGs co-circulating over time. While we recovered complete gene sequences from a single VG at W0, W3 and W5, we recovered complete gene sequences from two different VG at W1, W4 and W8, and from all three VGs at W2, W6 and W7. Moreover, 53 samples (63.8%) contained eight complete IAV gene segments, 13 (15.7%) contained more than eight and 17 samples (20.5%) contained less than 8 gene segments. Furthermore, 70 samples (84.3%) contained only sequences from a single IAV VG although 13 (15.7%) had gene segments from more than one VG (VG1 and VG2, n=1; VG1 and VG3, n=10; or VG1, VG2, and VG3, n=2). Additionally, six samples contained two different IAV antigenic subtypes (H1 gamma plus H1 beta, n=1; H1 gamma plus H3 cluster IV, n=2; and N1 plus N2, n=5 (Fig 17)).

Out of all complete IAV gene segments obtained (n=649), 78 (12%) and 83 (12.7%) were HA and NA sequences, respectively. The HA and NA pairwise sequence identity is illustrated in Fig 18. HA VG1 (subtype H1 gamma) sequences (n=48) were assembled between W0 and W7 and their percent sequence identity ranged between 98.2 and 100%. In contrast, HA VG2 (subtype H1 beta) sequences were only assembled at W4 and W14 and were 100% identical. Additionally, HA VG3 (subtype H3 cluster IV) sequences were identified

only between W6 and W8 and their percent sequence identity ranged between 99.9% and 100%. Moreover, NA sequences (n=83) included 50 sequences from VG1 (subtype N1), two from VG2 (subtype N1) and 31 from VG3 (subtype N2). NA sequences from VG1 were found between W0 and W7 while NA sequences from VG3 were assembled only at W1, W6, W7 and W8. The pairwise percent identity among NA sequences ranged between 99.6% and 100% for VG1 and between 98.8 % and 100% for VG3. In contrast, NA sequences from VG2 (n=2) were only assembled from W4 and W14 and were 100% identical. The consensus sequences of each internal gene segment (genes 1,2,3,5,7 and 8) were aligned using ClustalW and also showed a clear distinction between VG1, VG2, and VG3 for all internal gene segments (Fig 19). Within VGs (VG1, VG2 or VG3) the lowest percent identity among all gene segments was observed for gene segment 7 (matrix) and 8 (non-structural) (Fig 19e and 19f respectively).

The hypothetical HA and NA proteins for VG1, VG2 and VG3 were deduced to visualize differences within viral groups. For VG2, the two HA and NA sequences were 100% identical; thus, no further protein predictions were performed. In contrast, VG1 and VG3 had different HA and NA sequences, and some variants circulated more frequently than others (Fig 20). Nine different HA and nine NA amino acid sequences were translated from VG1 (Fig 20a and 20b respectively). Of 48 HA proteins from VG1, the majority (n=36) corresponded to two specific amino acid sequences, visualized in Fig 20a as node 1 (n=21) and node 2 (n=15); sequences represented in these two nodes differed by a single amino acid (T287A) and circulated until W6. The additional HA sequences (n=6) from VG1 were variants of the two major nodes. Only one HA variant was present at W0 for VG1 (Fig 20a) while three different NA proteins were found (Fig 20b). However, only the sequence represented in NA node one was identified in more than two weeks. Furthermore, six additional NA variants were identified after W0, but only those represented in nodes four and eight were identified from more than one sampling week (Fig 20b).

In contrast, only four different HA and four NA proteins were translated from VG3 (Fig 20c and 20d respectively). The most frequent HA proteins (node 1) from VG3 was first identified at W6 and persisted until W8 while the other three HA proteins were only identified at W7 and W8 (Fig 20c). The first NA protein from VG3 was identified at W1 and was not recovered again throughout the study period (Fig 20d). Furthermore, the most frequent NA variant of VG3 was identified at W6 and proceeded for two additional weeks while the remaining two variants were found only at W7 (Fig 20d).

To understand IAV infection and re-infection in pigs after weaning we compared the nucleotide and amino acid HA sequences of those samples sequenced from the same pig over time. Fifty-six samples collected from 26 pigs were used for this analysis (Fig 21). Five out of these 26 pigs (19.2%) had IAVs from the same VG (VG1 or VG3) and the HA pairwise nucleotide sequence identity within pig ranged between 98.2% and 100%. Additionally, 23 out of these 26 pigs (88.5%) met our definition of re-infected pigs (tested IAV positive in two non consecutive weeks). Complete HA sequences representing VG1 and VG3 were assembled from pig 41 at W6 and W7 (Fig 21); therefore, animal 41 was also considered a re-infected case for the sequencing analysis. Most IAV re-infections (19 out of 24, 79.2%) happened with IAVs from different VGs (VG1 and VG3, n=18; and VG1 and VG2, n=1). However, in five re-infected pigs (20.8%) the same VG (VG1) was recovered and the percent of sequence identity within pig for these viruses ranged between 99.6 and 99.9%. Moreover, when we compared the hypothetical amino acid sequences within pigs, we found that 2 out of 5 pigs (40%) infected with IAV in consecutive weeks had HA sequences that differ in 1 or 10 amino acids respectively while the other 3 pigs (60%) had the same HA protein (100% identical). In contrast, all 5 pigs re-infected with the same VG had HA proteins that were different in 1 or 2 amino acids (Fig 21).

Discussion

To understand the long-term persistence of influenza A viruses (IAVs) in pigs after weaning we designed a prospective cohort study and followed 132 3-week old pigs for 15 weeks after weaning. Using molecular detection, deep genome sequencing and hypothetical predictions of the main antigenic proteins of IAVs we demonstrated the complexity of IAV epidemiology and molecular diversity during infection of pigs under field conditions. These findings are important because they contribute to the understanding of persistence of IAVs in swine populations and should help us design better health interventions to prevent and control the disease in pigs. Additionally, prospective cohort studies in pigs could serve as an excellent animal model to study IAV transmission and evolution in humans, or to test the efficacy of health interventions given the physiological and immunological similarities between humans and pigs (186), the homology between human and swine IAVs, and the similarity on IAV transmission among humans and pigs (1, 24, 76).

We found that at the population level the persistence of IAV infection was the result of two contiguous IAV epidemics in which three distinct viral groups (VGs) co-circulated. Each epidemic wave of IAV infection was dominated by VG1 and VG3 respectively. However, VG1 and VG3 were both found co-circulating throughout the study period. Furthermore, a third virus group (VG2) was found co-circulating with VG1 and VG3 without dominating in any of the epidemic waves identified. Hence we determined that the endemic infection of IAV in this population was due to multiple VG causing contiguous epidemic waves. A similar epidemic trend of swine IAV infection has been described before in pigs after weaning in Europe (187) and might explain why swine populations are continuously infected with IAVs (17, 37, 128). In the Midwestern USA, 90% of pig-production herds (with growing pigs) are considered positive to IAVs (17) and in Europe swine IAVs are also widely distributed in pig farms (18). To our knowledge this is the first prospective cohort study that evaluated in detail the

complete genome of IAVs during transmission under field conditions and demonstrated that at the population level swine IAVs are also transmitted over time as a “cloud” of genotypes that are either closely related to each other or clearly distinct (different subtypes).

The co-circulation of different IAV alleles (sequence variants) during infection of children (169), horses (126), pigs (73, 141), and dogs (5) has been demonstrated before. The plasticity of swine IAV at the individual level (73, 141) might translate into our findings at the population level. This plasticity of the IAV genome over time may play a key role in the continuous outbreaks of IAVs and IAV re-infection of pigs after weaning. RNA viruses have higher mutation rates compared to many other microorganisms in nature (44, 188). Hence, finding a “cloud” of IAVs closely related to each other at the population level was expected. However, our findings unraveled a deeper layer of IAV diversity during infection of pigs because they demonstrated that this “cloud” of genotypes (virus groups) could exist concurrently for different VGs changing dynamically over time.

It is unclear in this study why the different VGs presented themselves at different times throughout the study. Different IAVs are expected to circulate in pigs after weaning (17, 18). However, different IAVs replicate at different rates under different environments (70, 71, 171, 175), which can affect viral fitness and transmission rates (189). For example in pigs, maternally derived antibodies (38) and acquired immunity (64) affect the transmission rates of IAVs among weaned piglets. If two different IAVs are transmitted at different rates then the proportion of IAVs found over time could change as observed in this study. The effect of virus fitness, transmission rates, and host immunity on IAV diversity over time should be further investigated and could highlight the targets for future effective health intervention methods.

Moreover, we found the complete genome of VG2 (100% identical) at W4 and W14 and strong evidence of VG2 co-circulating with VG1 and VG3 at other

sampling weeks. This finding indicated that some IAV genotypes might hiddenly replicate underneath the overall cloud of IAV genotypes and could only be recovered when the epidemic of the dominant VGs have ended. A recent study in North America illustrated how new genetic lineages of swine IAVs have emerged without being noticed for several years (76) while others that were commonly found before (e.g. H1 beta IAVs) now circulate unnoticed (75, 76). Therefore we speculate that the diversity and evolution of swine IAVs might be greatly underestimated or biased towards those genotypes that are over-represented during IAV epidemics or that are selected by procedures of virus culture (171). Our approach to evaluate the complete genome of IAVs directly from the nasal swab samples using NGS technologies should provide a greater resolution of IAV variants circulating in pig populations during infection and should help provide a better representation of the viruses that are circulating.

In this study we found two epidemic waves of IAV infection that overlapped at some point between W3 and W6 and could have allowed different VG to exchange gene segments. However most samples contained consensus genome constellations that represented a single VG although several samples contained gene segments from different IAVs. The segmented genome of IAV allows two or more viruses to exchange gene segments during infection (genetic reassortment) (9, 41) which is an important mechanism of virus diversification and emergence of novel IAV with zoonotic (11, 144) and pandemic potential (10). In pigs the antigenic genes are swapped between IAVs at different rates (72) and indicates that IAV reassortment in pigs is not a random event. However it is not know in what phase of IAV epidemics reassortment happens. If a single VG is dominating an epidemic reassortment might be less likely to happen. However, if two epidemics overlap reassortment might take place. In other species, the frequency of genetic reassortment among IAVs closely related to each other is random (70) while viruses evolving from very distinct genetic lineages are more restrictive (71). Therefore, further investigation is required to understand swine IAV reassortment events since this mechanism of virus diversification has changed

the genetic makeup of the virus during the past three decades (16, 54) and allowed the emergence of the first human pandemic of the 21st century (53).

It is assumed that most pigs at weaning would have some sort of MDA to different IAVs. Hence it is expected the HA and NA will be under higher immune pressure during infection after weaning compared to other viral gene segments because HA and NA are the main antigenic proteins of the virus. However, the highest genetic diversity within IAV gene segments from the same VG was found for gene segment seven (matrix) and eight (non-structural) and not for segments four (HA) and six (NA). Furthermore, the distance matrix for the percent pairwise sequence identity of all IAV gene segments allowed us to differentiate the same IAV viral groups. One previous study showed that all IAV gene segments, and not only HA and NA, present a dynamic distribution during infection of pigs with immunity to different IAVs (141). Another study, demonstrated that the substitution rate for HA1 was not different between pigs with or without active immunity to IAVs (73). Although, immune selection could depend on the affinity of antibodies to a given virus, IAV antigenic selection due to existing immunity is poorly understood and should be further investigated, especially to understand IAV immune evasion and re-infection in pigs after weaning.

During this cohort study, most cases of IAV re-infection happened with IAV containing different antigenic subtypes (H1N1 and H3N2), which is expected given the antigenic differences between H1 and H3 IAVs. However, in this study re-infection also happened with viruses that differed at the HA level in only one or two amino acids. Whether these amino acid differences changed or not the antigenic properties of the virus allowing IAV re-infection to happen needs to be further investigated in future studies. IAV re-infections have been described and characterized in humans (190-192). In pigs the mechanisms that allow IAVs to re-infect an animal are not clearly defined and need further investigation. Multiple factors have been proposed to explain IAV re-infection which include host, environment and virological factors (193) such as antigenic drift (63), differences

in cross protection among IAVs phenotypes (38, 194), immune response (142, 190), maternally derived immunity (143, 195) and the competition between naïve and memory B cells (196). One single amino acid difference within HA can change the antigenicity and receptor binding avidity of IAVs in humans and other animal species (63, 197, 198) and several polymorphic amino acid sites have been associated with antigenic differences among swine H1 viruses (56, 76), and 6 to 7 amino acid sites appear to be very important among swine and human H3 IAVs (197, 199).

This study also provides additional information in regards to epidemiological findings relevant to what happens within swine populations. The number of prevalent cases between each epidemic peak of IAV infection at week 2 (W2) and 7 (W7) respectively was not statistically different suggesting that the threshold point at which swine IAV epidemics subside was similar and that the transmission pattern of two different IAVs (VG1 and VG3) within the same population followed the same trend. This information is valuable because if transmission patterns of IAV infection in pigs are predictable then health interventions could be design to prevent them. Moreover, the prevalence and incidence density of IAV was different between males and females at some weeks after weaning. Unfortunately, males and females were housed in two different sides of the barn (right and left side respectively) and were separated by a hallway that did not allow nose-to-nose contact between pigs of different sex. Therefore, it was not possible to differentiate if these statistical differences were associated to the pigs' sex or their location within the pig barn. The association between influenza A virus and sex is not clear in humans (200). While some studies have shown that there are differences in viral infection patterns between males and females not associated to specific behaviors (200, 201), others have shown that hospitalization rates in children associated with respiratory infections can be different between males and females (202). However there is no information available regarding this issue in swine populations. Nevertheless, if sex is not likely to be associated with swine IAV infection then our results

indicated that some pens were more likely to test positive than others over time, which could be an indication of how the virus was transmitting between pens during the study period.

We recognize the limitations of our study design. Our results might only represent the population of pigs from which the cohort was selected. However, the pig farm selected for this study represented a common commercial wean-to-finish farm in the USA and the epidemiological findings of this study are similar to prior studies in the USA (17, 37), Europe (187), and field reports by veterinarians in the USA. Additionally, our random sample selection and large sample size minimized the random error and allowed us to make conclusions at the population level. Moreover, the detection of IAV by RRT-PCR might not represent true infections, as it cannot differentiate the stage of IAV infection or whether replicating virus was present, which needs to be taken into consideration for all data interpretation. Additionally, it is possible that the estimated genetic diversity of IAVs during this study is biased because we only selected a set of positive samples. However, we sequenced 25% of all positive samples and performed the genome amplification directly from nasal swabs, which avoided genotype selection during IAV isolation. The number of samples tested and sequenced in this study is probably the largest in any cohort study done to date in pigs after weaning.

In conclusion we demonstrated the complexity of IAV infection and re-infection in pigs after weaning in a large population of pigs under field conditions. We illustrated the dynamic diversity of the complete genome of IAVs over time and characterized its effect on the main antigenic proteins of the virus (HA) and (NA). The prolonged persistence of IAVs in pigs after weaning can be the result of multiple IAV epidemics that take place repeatedly over time or the re-infection with IAVs that are closely related to each other. These findings are important for decision making to control IAVs in pigs after weaning and to better understand virus diversity and emergence of IAVs in endemically infected swine populations.

Table 23. Epidemiological findings.

Number of pigs that tested influenza A virus (IAV) positive by real time reverse transcriptase polymerase chain reaction (RRT-PCR) distributed by week (W) and sex. Prevalent and incident cases (incidence density) are shown by week and compared between males and females.

W	Total number of pigs	Males*	Females	Total prevalent cases (rate per 100)	Prevalent cases among males (rate per 100)	Prevalent cases among females (rate per 100)	Incident cases (cases per 100 pigs-week)	Incident cases among males (cases per 100 males-week)	Incident cases among females (cases per 100 females-week)	Incidence density ratio between males and females (95% CI)
0	132	77	55	27 (20.5)	17 (22.1)	10 (18.2)	NA	NA	-	
1	132	77	55	70 (53.0)**	47(61.0)	23 (41.8)	49 (46.7)	32 (53.3)	17 (37.8)	1.4 (0.8,2.5)
2	131	77	55	86 (65.2)**	42 (54.5)	44 (80.0)	44 (71.0)	17 (56.7)	27 (84.4)	0.7 (0.4,1.2)
3	131	77	55	10 (7.6)	6 (7.8)	4 (7.3)	4 (8.7)	4 (11.4)	0 (0)	NA
4	131	76	55	5 (3.8)	3 (3.9)	2 (3.6)	4 (3.3)	2 (2.8)	2 (3.9)	0.7 (0.1,5.1)
5	131	76	55	6 (4.6)	5 (6.6)	1 (1.8)	6 (4.8)	5 (6.8)	1 (1.9)	3.6 (0.5,27)
6	130	76	55	21 (16.0)**	17 (22.4)	4 (7.3)	20 (16.0)	16 (22.5)	4 (7.4)	3.0 (1.1,8.6)**
7	130	76	55	76 (58.0)	43 (56.6)	33 (60.0)	66 (60.0)	35 (59.4)	31 (60.8)	1.0 (1.6,0.6)
8	129	75	55	38 (29.2)	19 (25.3)	19 (34.5)	17 (30.9)	10 (30.3)	7 (31.8)	1.0 (2.5, 0.4)
9	129	74	55	7 (5.4)	4 (5.4)	3 (5.5)	3 (3.3)	3 (5.4)	0 (0)	NA
10	129	74	55	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	NA
11	128	74	55	1 (0.8)	1 (1.4)	0 (0)	1 (0.8)	1 (1.4)	0 (0)	NA
12	128	73	55	5 (3.9)	4 (5.5)	1 (1.8)	5 (3.9)	4 (5.5)	1 (1.8)	3.0 (0.4,24,2)
13	128	73	55	5 (3.9)	1 (1.4)	4 (7.3)	5 (4.1)	1 (1.4)	4 (7.4)	0.2 (1.4,0.0)
14	127	73	55	5 (3.9)	4 (5.5)	1 (1.8)	5 (4.1)	4 (5.6)	1 (2.0)	2.8 (0.3,23.0)
15	127	72	55	5 (5.5)	4 (8.3)	1 (1.8)	7 (5.7)	6 (8.7)	1 (1.9)	4.7 (0.7,32.0)

* Males and females were located on the right and left side of the barn respectively

**Indicate significant difference ($p<0.05$) between males and females.

Table 24. Pairwise sequence comparison among influenza A virus gene templates.

Percent identity (ClustalW alignment) between gene segments of three representative influenza A viruses from virus group 1 (VG1, H1 gamma), 2 (VG2, H1 beta), and 3 (VG3, H3 cluster IV) that were used to map all Illumina sequencing reads.

Segment	VG1 vs. VG2	VG1 vs. VG3	VG2 vs. VG3
1 (PB2)	95.4	94.9	95.7
2 (PB1)	95.4	95.7	95.6
3 (PA)	94.8	91.8	95.5
4 (HA)	90.7	51.8	52.4
5 (NP)	94.3	93.1	96.8
6 (NA)	95.5	51.3	51.8
7 (M)	88.5	88.1	98.0
8 (NS)	95.9	95.6	96.6

Table 25. Frequency distribution of complete influenza A virus sequences by virus group (VG1, H1 gamma; VG2, H1 beta; and VG3, H3 cluster IV) and gene segment.

Polymerase base 2 (PB2), polymerase B1 (PB1), polymerase A (PA), hemagglutinin (HA), neuraminidase (NA), matrix (M), and non-structural protein (NS).

Virus	Gene segments								Total
	PB2 (%)	PB1 (%)	PA (%)	HA (%)	NP (%)	NA (%)	M (%)	NS (%)	
VG1 - H1 gamma	46 (60.5)	43 (59.7)	48 (64.0)	48 (61.5)	49 (62.0)	50 (60.2)	59 (63.4)	58 (62.4)	402 (61.9)
VG2 - H1 beta	2 (2.6)	2 (2.8)	2 (2.7)	2 (2.6)	2 (2.5)	2 (2.4)	4 (4.3)	4 (4.3)	19 (2.9)
VG3 - H3 cluster IV	28 (36.8)	27 37.5)	25 (33.3)	28 (35.9)	28 (35.4)	31 (37.3)	30 (32.3)	31 (33.3)	228 (35.1)
Total	76 (100)	72 (100)	75 (100)	78 (100)	79 (100)	83 (100)	93 (100)	93 (100)	649 (100)

Figure 14. Number of pigs and sampling times that pigs tested positive to influenza A viruses.

Each bar represents the number of pigs with 0 to 6 IAV positive swabs during the study period.

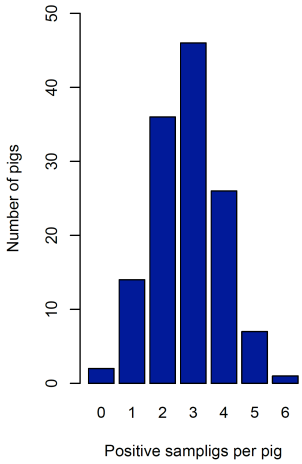


Figure 15. Prevalent and incident cases of influenza A virus distributed by week.

Green bars represent the total number of RRT-PCR positive cases per week (prevalent cases) and yellow bars illustrate the number of new cases found every week (incident cases).

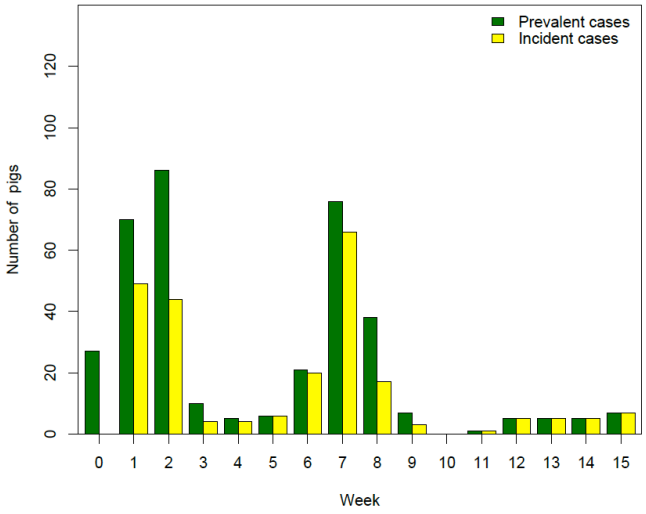


Figure 16. Distribution of Illumina sequencing reads for each gene segment over time.

Percentage of illumina sequencing reads mapped to influenza A virus gene segments (1 to 8) distributed by sample, week and viral groups one (VG1), two (VG2), and three (VG3). Each bar represents a sample. The y-axis indicates the percentage of reads by sample that were mapped to VG1 (H1 gamma, blue), VG2 (H1-beta, yellow) and VG3 (H3-cluster-IV, green) templates. The x-axis indicates the week from which each sample was selected.

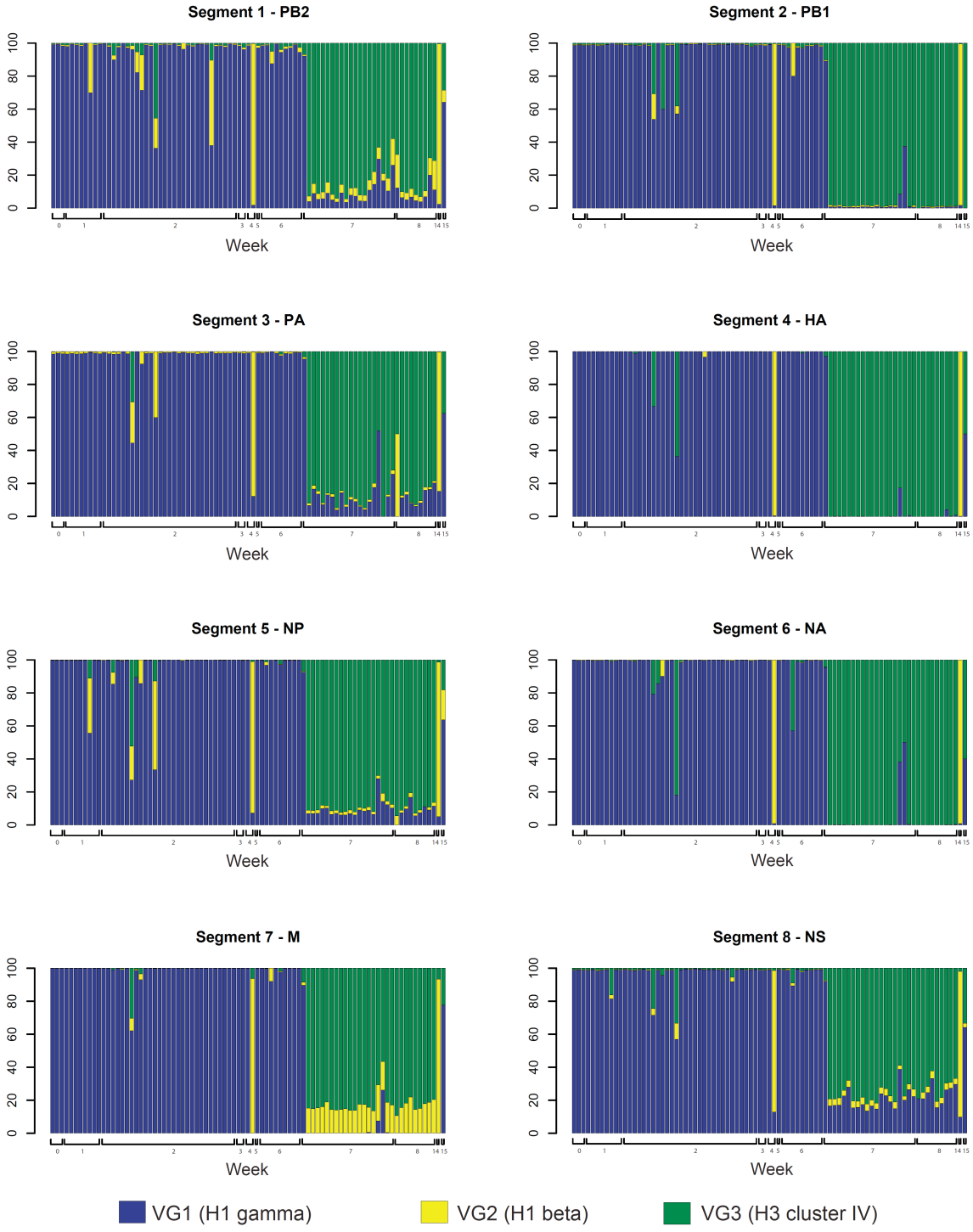
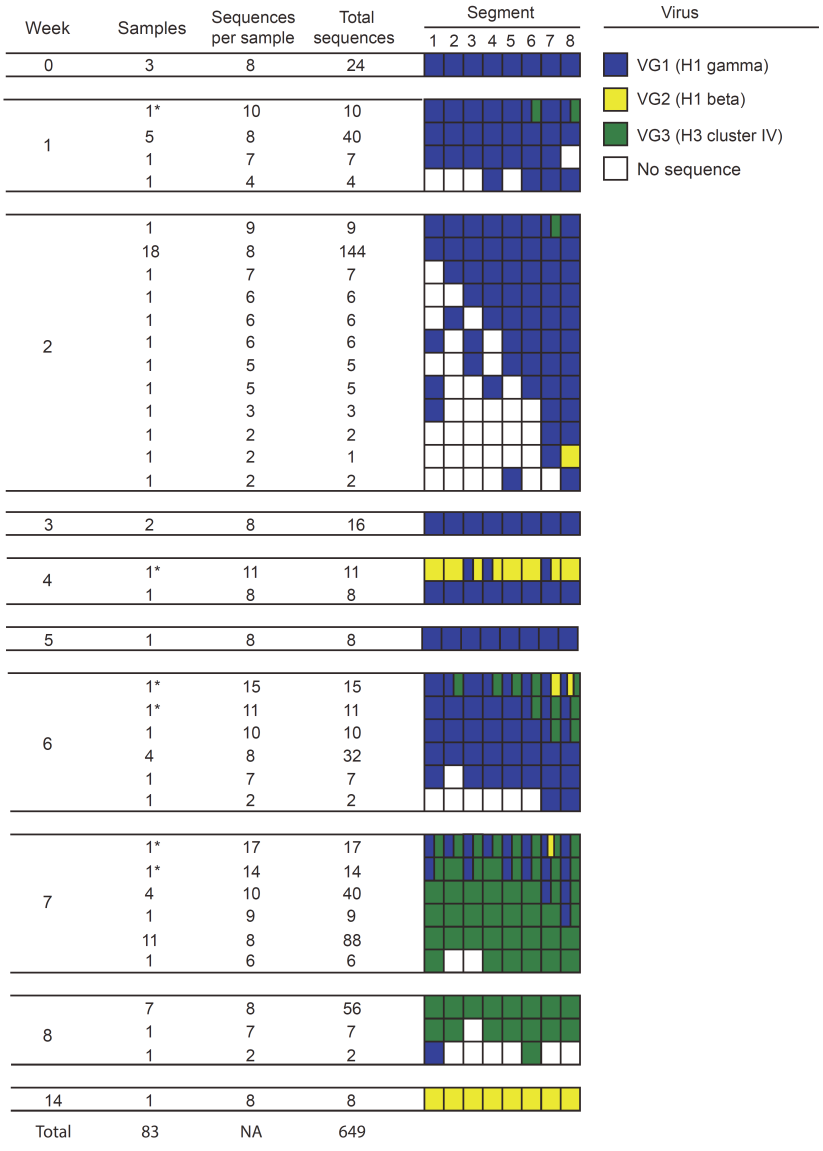

Segment 1 - PB2
Segment 2 - PB1
Segment 3 - PA
Segment 4 - HA
Segment 5 - NP
Segment 6 - NA
Segment 7 - M
Segment 8 - NS
0
20
40
60
80
100
Week
VG1 (H1 gamma)
VG2 (H1 beta)
VG3 (H3 cluster IV)

Figure 17. Influenza A virus (IAV) gene constellations distributed by week and sample.

The first four columns indicate the week (W0 to W14), number of samples sequenced (n=83), number of complete IAV gene segments per sample, and total number of sequences obtained. The remaining columns indicate IAV gene segments assembled based on viral group one (VG1, blue), two (VG2, yellow), and three (VG3, green). White boxes indicate that it was not possible to assemble a complete IAV gene segment.

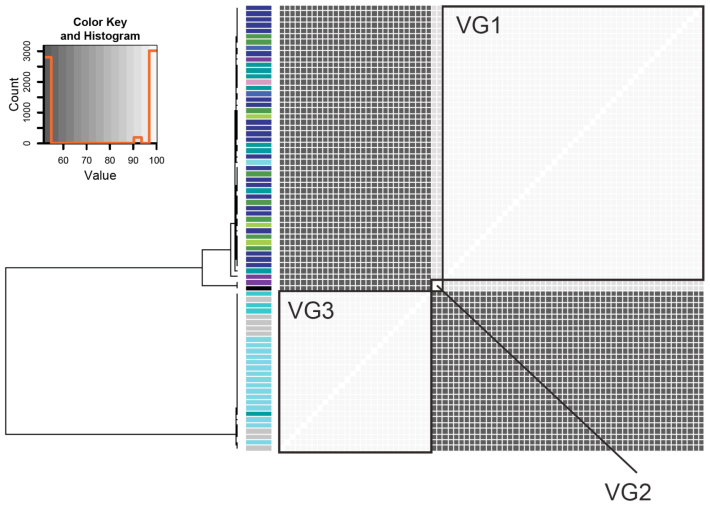


* Indicates samples with more than one antigenic subtype

Figure 18. Hemagglutinin (HA) and neuraminidase (NA) pairwise sequence identity.

Each heat map illustrates the percent sequence identity (ClustalW) among HA (a) and NA (b) sequences. The color key and histogram (x axis: percent identity; y axis: count) for each distance matrix is illustrated at the top left of each plot. Pairwise comparisons among virus group one (VG1, H1 gamma), two (VG2, H1 beta) and three (VG3, H3 cluster IV) are highlighted within black boxes. Dendrograms are distance based and illustrate the phylogenetic relationships between sequences. The bar side color at the left of each heat map indicates the sampling week (week 0 (W0) to week 14 (W14)) for each sample. The HA and NA pairwise comparison within and between groups is shown at the bottom of each plot.

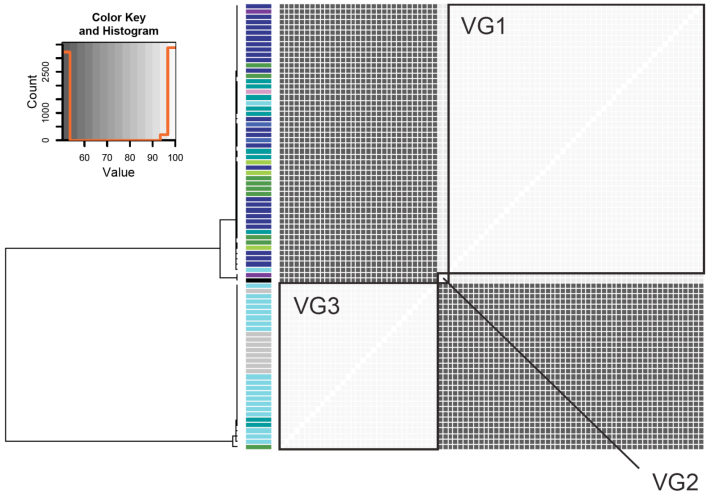
a. Segment 4 - Hemagglutinin (n=78)



Nucleotide percent identity HA

Virus group	1	2	3
1 (n=48)	98.2 - 100		
2 (n=2)	90.7 - 90.8	100	
3 (n=28)	51.6 - 52.0	52.1 - 52.2	99.9 - 100

b. Segment 6 - Neuraminidase (n=83)



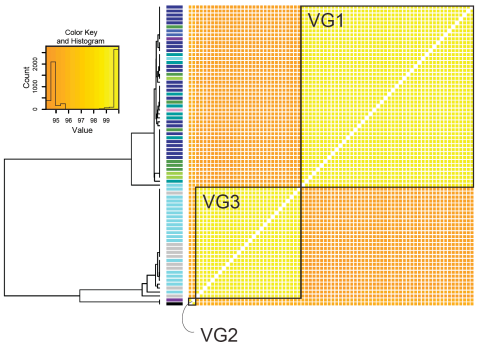
Nucleotide percent identity NA

Virus group	1	2	3
1 (n=50)	99.6 - 100		
2 (n=2)	95.4 - 95.6	100	
3 ((n=31)	50.3 - 50.9	51 - 51.4	98.8 - 100

Figure 19. Pairwise sequence identity (ClustalW) among influenza A virus (IAV) internal genes.

The color key and histogram (x axis: percent identity; y axis: count) for each distance matrix is illustrated at the top left of each plot. Pairwise comparisons among virus group one (VG1), two (VG2), and three (VG3) are highlighted within black boxes. Dendrograms are distance based and illustrate the phylogenetic relationships between sequences. The bar side color at the left of each heat map indicates the sampling week for each sample (week 0 (W0) to week 14 (W14)). The pairwise comparison within and between virus groups is shown at the bottom of each plot.

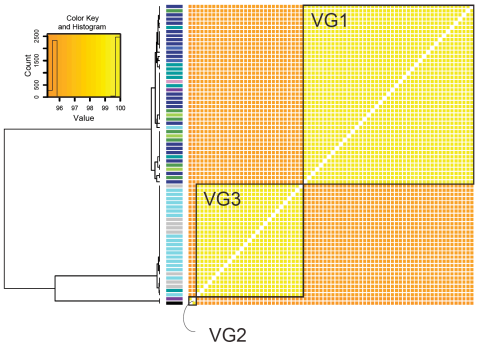
a. Segment 1 - Polymerase B2 (n=76)



Nucleotide percent identity PB2

Virus group	1	2	3
1 (n=46)	98.8 - 100		
2 (n=2)	95.2 - 95.4	100	
3 (n=28)	94.2 - 95.6	95.2 - 95.7	99.2 -100

b. Segment 2 - Polymerase B1 (n=72)

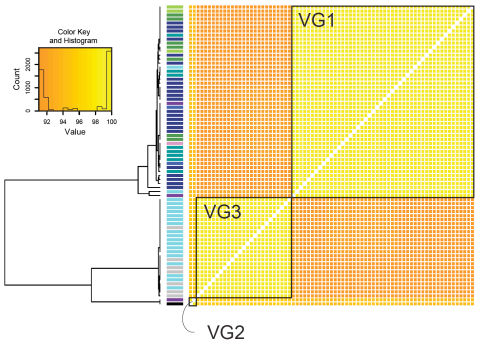


Nucleotide percent identity PB1

Virus group	1	2	3
1 (n=43)	99.6 -100		
2 (n=2)	95.2 - 95.3	100	
3 (n=27)	95.6 - 95.7	95.4	99.9 - 100

Week
0
1
2
3
4
5
6
7
8
14

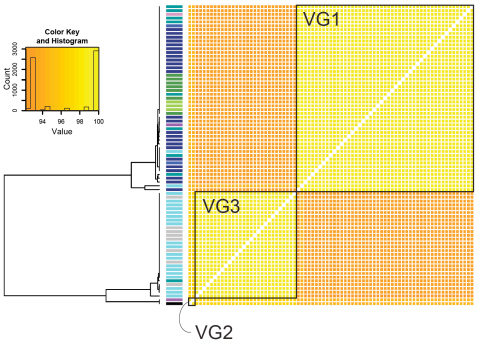
c. Segment 3 - Polymerase A (n=75)



Nucleotide percent identity PA

Virus group	1	2	3
1 (n=48)	97.8 - 100		
2 (n=2)	94.0 -95. 0	100	
3 (n=25)	91.0 - 92.8	95.4 - 95.5	99.9 - 100

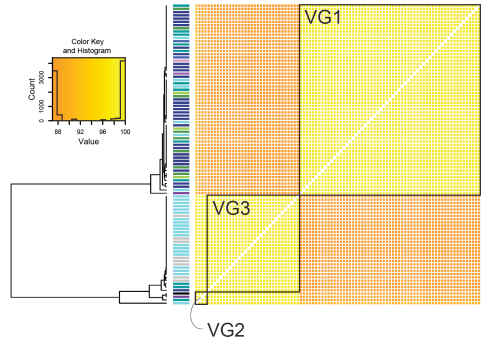
d. Segment 5 - Nucleoprotein (n=79)



Nucleotide percent identity NP

Virus group	1	2	3
1 (n=49)	97.7 - 100		
2 (n=2)	94.1 - 94.7	100	
3 (n=28)	92.2 - 94.3	96.4 - 96.7	99.9 - 100

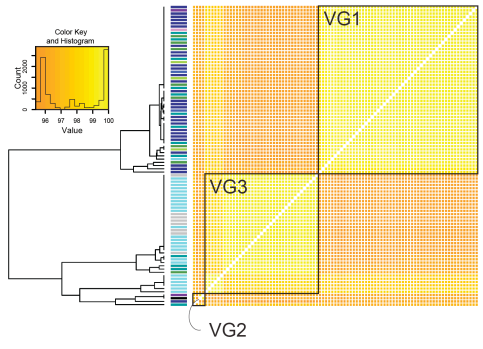
e. Segment 7 - Matrix (n=93)



Nucleotide percent identity M

Virus group	1	2	3
1 (n=59)	98.1 - 100		
2 (n=4)	87.7 - 91.4	96.3 - 100	
3 (n=30)	87.0 - 88.1	97.8 - 98.3	96.3 - 100

f. Segment 8 - Non-structural (n=93)

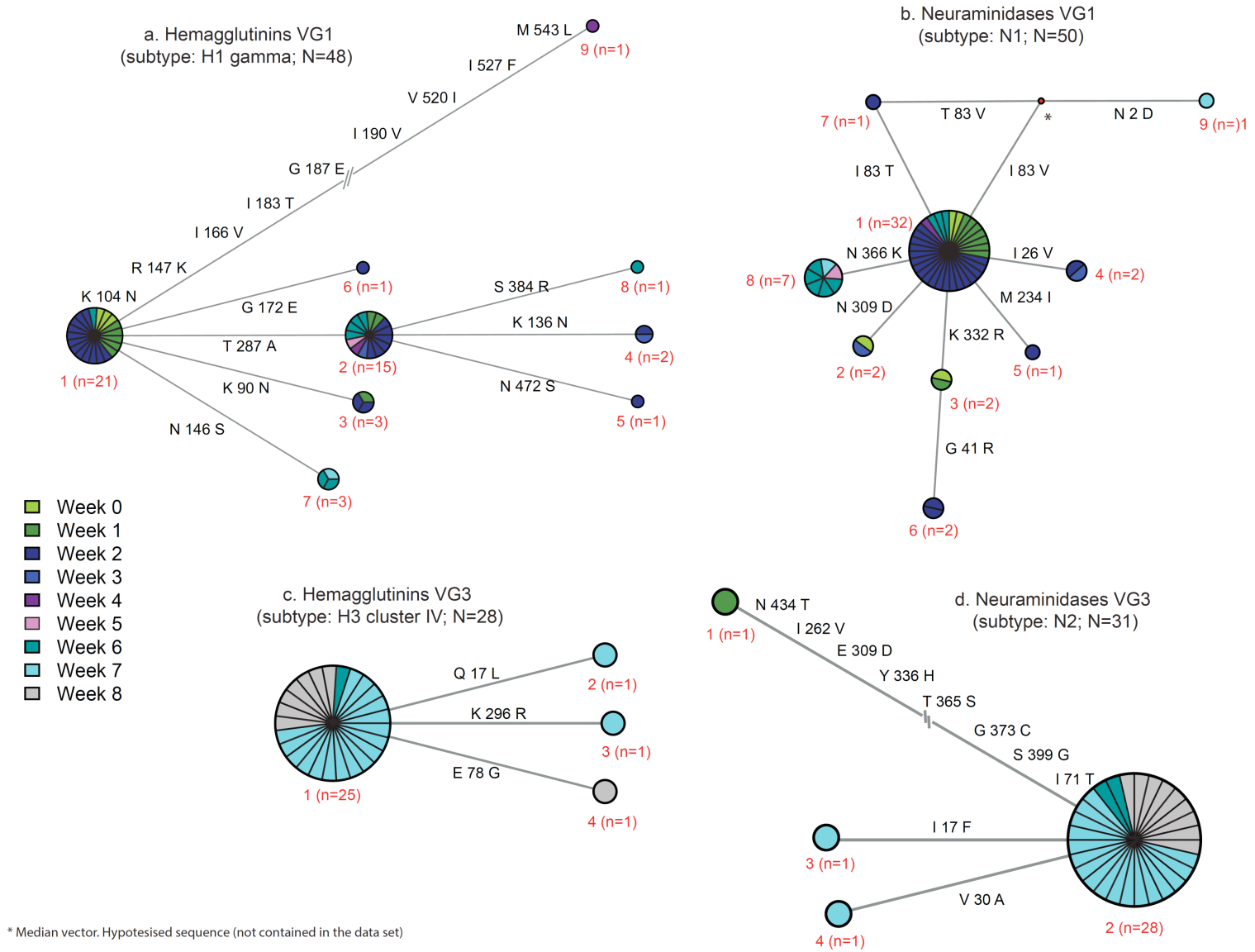


Nucleotide percent identity NS

Virus group	1	2	3
1 (n=58)	98.2 - 100		
2 (n=4)	95.8 - 97.0	95.5 - 100	
3 (n=31)	95.6 - 98.6	95.8 - 97.6	95.4 - 100

Figure 20. Network analysis of hemagglutinin (HA) and neuraminidase (NA) protein sequences of virus group one (VG1, H1 gamma) and three (VG3, H3 cluster IV).

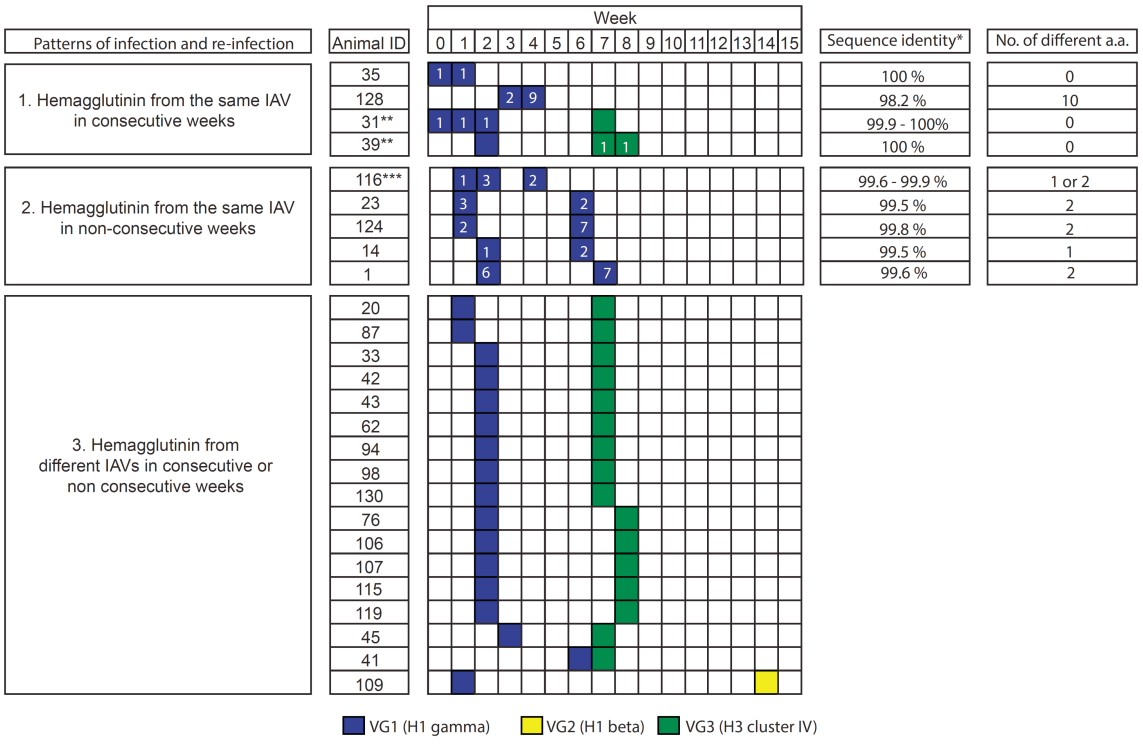
Network plots represent the relationships among HA and NA protein sequences for VG1 (panels a and b) and VG3 (panels c and d). Each node (circles) represents a protein sequence and its size is proportional to the number of sequences per node. Red numbers indicate the node number and sequence frequency (n) per node. Furthermore, nodes are color-coded according to sampling weeks (W0 to W8). The distance between nodes is proportional to the number of amino acid differences between sequences (larger distances are indicated as “-//-“ to fit the plot) and polymorphic sites (position and amino acid variants) are indicated between nodes.



* Median vector. Hypotesised sequence (not contained in the data set)

Figure 21. Patterns of influenza A virus (IAV) infection and re-infection.

Fifty-six hemagglutinin (HA) nucleotide sequences obtained from 26 pigs over time (week 0 to week 15) are illustrated based on virus group one (VG1, H1 gamma blue), two (VG2, H1 beta yellow), and three (VG3, H3 cluster IV green). Three different patterns of IAVs are shown: 1) HA sequences from the same IAV group found within a pig in consecutive weeks; 2) HA sequences from the same IAV group found within a pig in non-consecutive weeks; and 3) HA sequences from different IAVs found within a pig in consecutive or non-consecutive weeks. White numbers within colored boxes indicate the HA node in which this sequence clustered at the amino acid level as indicated in Fig 20.



* Percentage of nucleotide sequence identity between sequences from the same virus group (VG)
** Animals 31 and 39 were positive to the same VG (VG1 and VG3 respectively) in consecutive weeks and to VG1 and VG3 in non consecutive weeks.
*** Animal 116 was positive for the same VG (VG1) in consecutive and non-consecutive weeks.

Chapter 7: General discussion and conclusions

Influenza A viruses (IAV) are distributed worldwide and are able to infect many animal species including humans and pigs (9). A reassorted IAV caused the 2009 IAV pandemic (10), highlighted the public health risk of swine IAVs, and boosted the research on the molecular evolution of swine IAVs. Hence, the understanding of swine IAV evolution at the regional and global scales has improved significantly since 2009. In pigs, IAV is a major cause of respiratory disease and is associated with substantial production losses (16, 33, 34). However there is limited information available on the epidemiology of IAV in swine and how the genetic diversity of the virus affects IAV epidemiology in pigs, which is crucial to design better health interventions to control the disease in pigs and to reduce the risk to public health.

This PhD dissertation aimed at narrowing the gap of knowledge between aspects of the epidemiology of swine IAVs and how viral genetic diversity affects infection of pig populations. We integrated basic concepts of epidemiologic methods and contemporary techniques for virus sequencing and bioinformatics to characterize IAVs under field and experimental conditions. We focused these studies first on commercial breeding herds to evaluate the role of specific pig subpopulations on the epidemiology of IAV within and between herds (chapter 2) and characterized the complete genome of the virus over time in these subpopulations (chapter 3). Then we studied the antigenic drift of IAVs in 3-week-old pigs with or without MDA (chapter 4), estimated the diversity of the whole genome of IAV during infection of vaccinated pigs (chapter 5), and characterized the complete genome of several IAVs co-circulating naturally over time in a large population of pigs after weaning. At the herd level we also focused on the epidemiological characterization of IAV infections (chapter 2 and 6) and we analyzed the complete genome diversity of the virus to address basic questions on virus introduction, evolution, persistence and re-infection (chapter 3 and 6). At the individual animal level (chapters 4 and 5) we utilized experimental studies to

evaluate the effects of maternally derived antibodies (MDA) and active immunity (vaccination) on IAV diversity and antigenic drift. We focused our studies on evaluating the genetic diversity of the virus during transmission among weaned pigs with different immune statuses because these differences are common under field conditions and might lead IAVs into distinct evolutionary pathways (63, 67, 69, 203).

In chapter two, we confirmed the persistence of IAV over time in breeding herds and the co-circulation of more than one IAV subtype at a single sampling event. We proved that this long-term persistence of IAVs in breeding herds resulted in a higher odds of IAV infection in groups of new gilts (on farm for less than 4 weeks) and piglets compared to gilts (on farm for more than 4 weeks). Additionally, in these herds we found a strong association between IAV infection and year quarter indicating a seasonal pattern of IAV infection. We also showed that piglets and new gilts should be targets of IAV health interventions to reduce the impact of swine influenza on breeding herds and that there is value in having year around surveillance.

At the individual level IAV infections in pigs do not last more than one week (3, 16). However pig farms can test positive to IAV for prolonged periods of time although the mechanisms of virus persistence at the population level are not clearly understood. In the Midwestern USA, 90% of the farms with growing pigs are considered positive to IAVs and multiple IAVs are commonly found (17, 37). Moreover, the seasonality of IAV infections in pig herds in North America has been under debate given that IAVs can be recovered from pig farms year around (17, 75, 78). Our results confirmed that IAVs can be found for prolonged periods of time in swine breeding herds and supported the seasonality of IAV infection in pigs. Furthermore, this is the first study to take in consideration the roles that

specific animal subpopulations may play in the long-term persistence of IAVs in pig breeding herds.

Multiple pig subpopulations (e.g. different ages, production stages, times of arrival etc.) may perpetuate the maintenance of IAV over time because they can be at different risks of IAV infection and harbor different IAV strains. Furthermore, in pig breeding herds, gilts and piglets have high turnover rates and might represent a continuous niche for IAV replication. Piglets are born naïve to any IAVs (36) and with an unknown, but likely highly variable, component of MDA acquired after birth. Therefore they can potentially become infected with different IAVs present in the breeding herd. In addition, gilts with or without active immunity against specific IAVs can be susceptible to IAVs present in breeding herds and also represent of source for introducing new IAVs to herds. Therefore we believe health interventions to reduce IAV transmission in breeding herds should target piglets before weaning and gilts at arrival. However, for these interventions to be effective a detailed knowledge of IAV molecular diversity at the herd level is required.

In chapter 3, we used deep genome sequencing and hypothetical proteins to estimate the molecular diversity of those IAVs isolated from the breeding herds studied in chapter 2. We demonstrated that the continuous detection of IAV within pig farms could imply the presence of different genetic lineages and not necessarily the persistence of the same IAV over time. However, we also identified the same genetic lineage for prolonged periods of time. Furthermore, we found a dynamic distribution of IAV genotypes over time within and between pig subpopulations and documented the emergence, persistence, and subsidence of IAV genotypes. Studying the complete genome of IAVs isolated from pig subpopulations that are constantly fluctuating (i.e suckling piglets, new gilts, and gilts) increased our knowledge on the role of these populations on IAV

diversity. We demonstrated that different pig subpopulations within a farm could harbor different IAVs over time and that the co-circulation of multiple genotypes within a subpopulation could facilitate IAV reassortment. The complete genome characterization of IAVs using next generation sequencing technologies allowed us to differentiate IAVs over time with higher resolution than in any previous studies and unraveled a deeper layer of IAV genetic diversity during infection of pigs in breeding herds. We found multiple IAV genotypes that were either closely related to each other or clearly distinct. We also found several reassortant IAVs. Swine IAV reassortment is important because reassortment can result in dramatic changes in the antigenic properties of the virus (9, 41). However, it is not clear if this dynamic makeup of IAV genome in breeding herds, with long-term persistence of certain IAV genotypes is due to viral or host factors.

We hypothesize that there are three possible explanations for the persistence and genetic diversity of IAV in these pig breeding herds. First, there may be differences between IAVs from the same genetic lineage that enable the virus to evade herd immunity over time. Second, new gilts may be the source of the same IAV lineage and could introduce similar viral variants with each group introduced into the breeding herd; and third, the continuous availability of susceptible animals in breeding herds (newborns or new gilts) may allow the continuous replication of resident IAVs over time. Antigenic cartography (56, 76, 199) between those IAVs isolated from these herds could help to test the first hypothesis. Additionally, sampling gilts at arrival could help evaluate the role of new gilts on the introduction of new IAVs into breeding herds. However, IAV infection, replication and transmission will likely occur in the presence of enough susceptible individuals. Therefore, the continuous availability of susceptible individuals might be the most important cause of the persistence of IAVs at the population level in swine breeding herds.

When IAV positive pigs are weaned, they can serve as a source of IAVs to downstream swine sites and their respective regions (37). Therefore to better understand the molecular evolution of IAV during infection of weaned pigs we evaluated the genetic and antigenic diversity of the virus under different immune statuses in 3-week old weaned pigs. First we compared the antigenic differences at the HA level of IAVs among pigs with or without MDA against IAVs (chapter 4); subsequently we compared the complete genome plasticity of the virus in pigs with active immunity to IAVs (chapter 5); and finally we studied the transmission pattern and IAV genome diversity during infection of pigs after weaning under field conditions (chapter 6). In pigs with or without MDA against IAVs (Chapter 4) we found that nucleotide substitutions at the HA level can happen shortly after infection. Furthermore, we demonstrated that these nucleotide differences between HAs were not limited to the HA1 region of the HA but also happened within the signal peptide and the HA2 region. Our results are in agreement with a previous study that indicated that there was no difference between the evolutionary trends of IAV in pigs with or without immunity against IAVs (73). Furthermore, using deep genome sequencing we proved that the genetic diversity of IAVs during infection of weaned vaccinated pigs (chapter 5) is dynamic, within and between pigs, and not limited to the main antigenic genes of the virus. Our results also illustrated the importance of the internal gene segments on the assessment of IAV diversity during infection of pigs. The intra- and inter-host variability of the complete IAV genome during infection has been reported recently in children (169) and at the HA level in pigs (73), horses (4), and dogs (5). However, to our knowledge this is the first study reporting the complete genome characterization of IAV during infection of vaccinated pigs using deep genome sequencing.

MDA and humoral antibodies can change the transmission rate of IAV between pigs (38, 64) and variable transmission rates might influence genetic diversity

over time. However, the effects of humoral immune response on IAVs diversity may be associated with the cross reactivity of certain antibodies to different IAVs and their interference with an appropriate immune response (82, 142, 195, 204). Most weaned pigs in the contemporary swine industry are expected to have MDA against IAVs. Furthermore, pigs can be exposed to one or more IAVs over time during their lifetime and develop a variety of humoral immune response against IAVs (16, 18, 120). In our studies we did not find differences in the nucleotide substitutions that took place at the HA level in pigs with our without MDA and we found that all IAV gene segments had a dynamic genetic composition during infection of vaccinated pigs. However, different IAV strains or vaccines might lead to different results. The effect of humoral immunity on swine IAV evolution should be further investigated because it might shape the overall genetic diversity of the virus and could give better tools to design health interventions at the population level.

In chapter 6, we investigated in more detail the epidemiology of IAVs in pigs after weaning and attempted to identify molecular traits that might be associated with IAV re-infection. We confirmed that pigs can bring IAVs to other swine farms at weaning and that the long-term persistence of IAVs in pigs after weaning could be the result of different epidemic waves of IAV infection. Additionally, we found that different IAVs coexisted as a population of viruses that were either closely related to each other in the form of viral groups (VGs) or clearly distinct representing distinct IAV genetic lineages. These different VGs had a clear distribution over time with one different VG dominating each IAV epidemic wave identified during the study and illustrated the complexity of swine IAV genetic diversity. Furthermore, in this study we had convincing evidence that indicated that pigs could become re-infected with viruses that differed by as few as one amino acid in HA although most of IAV re-infections involved viruses from different genetic lineages. Understanding the molecular evolution of swine IAVs

under field conditions is important for vaccine selection, evaluation of IAV reassortment, and to determine the introduction or persistence of new IAVs. Additionally amplifying the complete genome directly from the nasal swab in some studies gave us a better understanding of viral diversity in the host and avoided bias due to selection during virus isolation in cell cultures.

Future studies are needed to test the efficacy of interventions targeting the aforementioned pig subpopulations. It is important to estimate the effect of animal movements within and between herds to determine the distribution and genetic diversity of IAVs. If animal movements are associated with the emergence and persistence of IAVs at the herd level then animal flows could be modified to reduce the risk of infection and maintenance. For example, sampling gilts at arrival, or introducing gilts at different ages or managing the flows all in/all out could help estimate the risk of new IAV introductions associated to new gilts arrival and their age at arrival. Additionally, at the herd level it will be important to characterize genetically and antigenically the persistent viruses in order to select the best vaccine available or to provide enough evidence that would require a policy change on swine IAV vaccine updates. Evaluation of vaccination protocols targeting timing, season, type of animals and type of vaccines (killed vs. live attenuated) are also warranted.

Finally, we recognize multiple limitations in our studies and they should be taken into account in the interpretation of our results. First at the population level the external validity of our results are limited because the number of pig farms studied was limited and farms were conveniently selected. Therefore, it is unclear how our results represent pig populations across US regions. Nevertheless all farms for our studies were selected to represent commercial pig production sites in the Midwestern USA. Additionally we were not able to estimate the association of IAV genetic traits and epidemiological findings because these types of studies

require larger sample sizes and that was out of the scope of this thesis. In addition, in order to fully investigate the effect of vaccination on virus evolution, it would be important to compare the variability of the IAV genome between pigs with or without active immunity against IAVs. In our studies in chapter 5, we did not have a negative control to estimate if the genetic differences found were due to IAV vaccination or happened because of other host or virus factors.

In conclusion we demonstrated that the complex dynamic of IAV diversity at the herd level is the result of the plasticity of IAV genome during infection of pigs regardless of their immune status. The plasticity of IAV genome during infection of pigs at the individual level indicated that there is a dynamic “cloud” of genotypes during virus replication that are closely related to each other, which might be translated at the population as different virus groups that can co-circulate with other influenza A virus groups over time. This variability on IAV genome at the individual and population levels during transmission of the virus under field conditions could affect virus fitness over time and allow the persistence of the virus in populations at higher risk of IAV infection. We found that swine IAV infections are not evenly distributed among all subpopulations present in pig breeding herds or in pigs after weaning and demonstrated that the long-term persistence of IAVs in pig farms could be associated with the continuous occurrence of IAV epidemics at the herd level. Hence, we hypothesize that the most important epidemiological factor for IAV persistence at the population level is the continuous availability of susceptible animals to IAV infections that allow this dynamic cloud of IAV genotypes to replicate over time. We suggest that health interventions to control IAV in swine populations and reduce its zoonotic potential should aim to reduce the transmission and persistence of IAVs among those pig subpopulations that are continuously introduced into pig farms (new gilts, newborn pigs, and weaned pigs). The work presented in this thesis contributes significantly to the understanding of IAV

diversity, persistence and evolution in pigs and provides useful information to mitigate IAV infections in pigs and its zoonotic potential.

References

1. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature. 2015;523(7559):217-U06.
2. Chen RB, Holmes EC. Hitchhiking and the Population Genetic Structure of Avian Influenza Virus. Journal of Molecular Evolution. 2010;70(1):98-105.
3. Torremorell M, Allerson M, Corzo C, Diaz A, Gramer M. Transmission of Influenza A Virus in Pigs. Transboundary and Emerging Diseases. 2012;59:68-84.
4. Murcia PR, Baillie GJ, Stack JC, Jervis C, Elton D, Mumford JA, et al. Evolution of equine influenza virus in vaccinated horses. J Virol. 87. United States2013. p. 4768-71.
5. Hoelzer K, Murcia PR, Baillie GJ, Wood JLN, Metzger SM, Osterrieder N, et al. Intrahost Evolutionary Dynamics of Canine Influenza Virus in Naive and Partially Immune Dogs. Journal of Virology. 2010;84(10):5329-35.
6. Ali A, Daniels JB, Zhang Y, Rodriguez-Palacios A, Hayes-Ozello K, Mathes L, et al. Pandemic and Seasonal Human Influenza Virus Infections in Domestic Cats: Prevalence, Association with Respiratory Disease, and Seasonality Patterns. Journal of Clinical Microbiology. 2011;49(12):4101-5.
7. Harder TC, Siebert U, Wohlsein P, Vahlenkamp T. Influenza A virus infections in marine mammals and terrestrial carnivores. Berliner Und Munchener Tierarztliche Wochenschrift. 2013;126(11-12):500-8.
8. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, et al. New world bats harbor diverse influenza a viruses. PLoS Pathog. 2013;9(10):e1003657.
9. Taubenberger JK, Kash JC. Influenza Virus Evolution, Host Adaptation, and Pandemic Formation. Cell Host & Microbe. 2010;7(6):440-51.
10. Rambaut A, Holmes E. The early molecular epidemiology of the swine-origin A/H1N1 human influenza pandemic. PLoS currents. 2009;1:RRN1003.
11. Bowman AS, Nelson SW, Page SL, Nolting JM, Killian ML, Sreevatsan S, et al. Swine-to-Human Transmission of Influenza A(H3N2) Virus at Agricultural Fairs, Ohio, USA, 2012. Emerging Infectious Diseases. 2014;20(9):1472-80.
12. Freidl GS, Meijer A, de Bruin E, de Nardi M, Munoz O, Capua I, et al. Influenza at the animal-human interface: a review of the literature for virological evidence of human infection with swine or avian influenza viruses other than A(H5N1). Eurosurveillance. 2014;19(18):8-26.
13. Nelson MI, Lemey P, Tan Y, Vincent A, Lam TTY, Detmer S, et al. Spatial Dynamics of Human-Origin H1 Influenza A Virus in North American Swine. Plos Pathogens. 2011;7(6).
14. York I, Donis RO. The 2009 Pandemic Influenza Virus: Where Did It Come from, Where Is It Now, and Where Is It Going? In: Richt JA, Webby RJ, editors. Swine Influenza. Current Topics in Microbiology and Immunology. 370. Berlin: Springer-Verlag Berlin; 2013. p. 241-57.

15. Nelson MI, Gramer MR, Vincent AL, Holmes EC. Global transmission of influenza viruses from humans to swine. Journal of General Virology. 2012;93:2195-203.
16. Vincent AL, Ma WJ, Lager KM, Janke BH, Richt JA. Swine Influenza Viruses: A North American Perspective. In: Maramorosch K, Shatkin A, Murphy F, editors. Advances in Virus Research, Vol 72. Advances in Virus Research. 722008. p. 127-54.
17. Corzo CA, Culhane M, Juleen K, Stigger-Rosser E, Ducatez MF, Webby RJ, et al. Active surveillance for influenza A virus among swine, midwestern United States, 2009-2011. Emerg Infect Dis. 2013;19(6):954-60.
18. Simon G, Larsen LE, Durrwald R, Foni E, Harder T, Van Reeth K, et al. European Surveillance Network for Influenza in Pigs: Surveillance Programs, Diagnostic Tools and Swine Influenza Virus Subtypes Identified in 14 European Countries from 2010 to 2013. Plos One. 2014;9(12):21.
19. Allerson MW, Cardona CJ, Torremorell M. Indirect Transmission of Influenza A Virus between Pig Populations under Two Different Biosecurity Settings. Plos One. 2013;8(6):9.
20. Corzo CA, Culhane M, Dee S, Morrison RB, Torremorell M. Airborne detection and quantification of swine influenza a virus in air samples collected inside, outside and downwind from swine barns. PLoS One. 2013;8(8):e71444.
21. Rajao DS, Anderson TK, Gauger PC, Vincent AL. Pathogenesis and Vaccination of Influenza A Virus in Swine. Influenza Pathogenesis and Control - Vol I. 2014;385:307-26.
22. Li WF, Shi WF, Qiao HJ, Ho SYW, Luo AR, Zhang YZ, et al. Positive selection on hemagglutinin and neuraminidase genes of H1N1 influenza viruses. Virology Journal. 2011;8.
23. Ducatez MF, Hause B, Stigger-Rosser E, Darnell D, Corzo C, Juleen K, et al. Multiple Reassortment between Pandemic (H1N1) 2009 and Endemic Influenza Viruses in Pigs, United States. Emerging Infectious Diseases. 2011;17(9):1624-9.
24. Nelson MI, Viboud C, Vincent AL, Culhane MR, Detmer SE, Wentworth DE, et al. Global migration of influenza A viruses in swine. Nat Commun. 62015. p. 6696.
25. Knauer MT, Hostetler CE. US swine industry productivity analysis, 2005 to 2010. Journal of Swine Health and Production. 2013;21(5):248-52.
26. Davies PR. Intensive Swine Production and Pork Safety. Foodborne Pathogens and Disease. 2011;8(2):189-201.
27. Oh SH, Whitley NC. Pork Production in China, Japan and South Korea. Asian-Australasian Journal of Animal Sciences. 2011;24(11):1629-36.
28. Nguyen TLT, Hermansen JE, Mogensen L. Fossil energy and GHG saving potentials of pig farming in the EU. Energy Policy. 2010;38(5):2561-71.
29. Batista L, Pijoan C, Torremorell M. Experimental injection of gilts with porcine reproductive and respiratory syndrome virus (PRRSV) during acclimatization. Journal of Swine Health and Production. 2002;10(4):147-50.
30. Torremorell M, Moore C, Christianson WT. Establishment of a herd negative for porcine reproductive and respiratory syndrome virus (PRRSV) from PRRSV positive sources. Journal of Swine Health and Production. 2002;10(4):153-60.

31. Linhares DCL, Cano JP, Torremorell M, Morrison RB. Comparison of time to PRRSv-stability and production losses between two exposure programs to control PRRSv in sow herds. Preventive Veterinary Medicine. 2014;116(1-2):111-9.
32. Fahrion AS, Beilage EG, Nathues H, Durra S, Doherr MG. Evaluating perspectives for PRRS virus elimination from pig dense areas with a risk factor based herd index. Preventive Veterinary Medicine. 2014;114(3-4):247-58.
33. Alvarez J, Sarradell J, Kerkaert B, Bandyopadhyay D, Torremorell M, Morrison R, et al. Association of the presence of influenza A virus and porcine reproductive and respiratory syndrome virus in sow farms with post-weaning mortality. Prev Vet Med. 2015;121(3-4):240-5.
34. Er C, Lium B, Tavornpanich S, Hofmo PO, Forberg H, Hauge AG, et al. Adverse effects of Influenza A(H1N1)pdm09 virus infection on growth performance of Norwegian pigs - a longitudinal study at a boar testing station. BMC Veterinary Research. 2014;10.
35. Kwit K, Pomorska-Mol M, Markowska-Daniel I. The influence of experimental infection of gilts with swine H1N2 influenza A virus during the second month of gestation on the course of pregnancy, reproduction parameters and clinical status. Bmc Veterinary Research. 2014;10.
36. Salmon H, Berri M, Gerdts V, Meurens F. Humoral and cellular factors of maternal immunity in swine. Developmental and Comparative Immunology. 2009;33(3):384-93.
37. Allerson MW, Davies PR, Gramer MR, Torremorell M. Infection Dynamics of Pandemic 2009 H1N1 Influenza Virus in a Two-Site Swine Herd. Transbound Emerg Dis. 2014;61(6):490-9.
38. Allerson M, Deen J, Detmer SE, Gramer MR, Joo HS, Romagosa A, et al. The impact of maternally derived immunity on influenza A virus transmission in neonatal pig populations. Vaccine. 2013;31(3):500-5.
39. Palase P. ***Orthomyxoviridae:*** **The Viruses and Their Replication**. In: Knipe D, Howley P, editors. Fields' virology. II. 5th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2007.
40. Ampofo WK, Baylor N, Cobey S, Cox NJ, Daves S, Edwards S, et al. Improving influenza vaccine virus selection: report of a WHO informal consultation held at WHO headquarters, Geneva, Switzerland, 14-16 June 2010. Influenza and other respiratory viruses. 2012;6(2):142.
41. Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. Nature. 2003;422(6930):428-33.
42. Latorre-Margalef N, Grosbois V, Wahlgren J, Munster VJ, Tolf C, Fouchier RAM, et al. Heterosubtypic Immunity to Influenza A Virus Infections in Mallards May Explain Existence of Multiple Virus Subtypes. Plos Pathogens. 2013;9(6):12.
43. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. Nature. 2008;453(7195):615-U2.

44. Borderia AV, Stapleford KA, Vignuzzi M. RNA virus population diversity: implications for inter-species transmission. Current Opinion in Virology. 2011;1(6):643-8.
45. Zhu X, Yu W, McBride R, Li Y, Chen LM, Donis RO, et al. Hemagglutinin homologue from H17N10 bat influenza virus exhibits divergent receptor-binding and pH-dependent fusion activities. Proc Natl Acad Sci U S A. 2013;110(4):1458-63.
46. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, et al. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science. 2004;303(5665):1838-42.
47. Kaplan BS, Webby RJ. The avian and mammalian host range of highly pathogenic avian H5N1 influenza. Virus Research. 2013;178(1):3-11.
48. Boivin S, Cusack S, Ruigrok RWH, Hart DJ. Influenza A Virus Polymerase: Structural Insights into Replication and Host Adaptation Mechanisms. Journal of Biological Chemistry. 2010;285(37):28411-7.
49. Hass J, Matuszewski S, Cieslik D, Haase M. The role of swine as "mixing vessel" for interspecies transmission of the influenza A subtype H1N1: A simultaneous Bayesian inference of phylogeny and ancestral hosts. Infection Genetics and Evolution. 2011;11(2):437-41.
50. Thacker E, Janke B. Swine influenza virus: Zoonotic potential and vaccination strategies for the control of avian and swine Influenzas. Journal of Infectious Diseases. 2008;197:S19-S24.
51. Khan SU, Atanasova KR, Krueger WS, Ramirez A, Gray GC. Epidemiology, geographical distribution, and economic consequences of swine zoonoses: a narrative review. Emerging Microbes & Infections. 2013;2.
52. Davies PR. One World, One Health: The Threat of Emerging Swine Diseases. A North American Perspective. Transboundary and Emerging Diseases. 2012;59:18-26.
53. Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459(7250):1122-U107.
54. Kitikoon P, Nelson MI, Killian ML, Anderson TK, Koster L, Culhane MR, et al. Genotype patterns of contemporary reassorted H3N2 virus in US swine. J Gen Virol. 2013;94(Pt 6):1236-41.
55. Lorusso A, Vincent AL, Gramer MR, Lager KM, Ciacci-Zanella JR. Contemporary Epidemiology of North American Lineage Triple Reassortant Influenza A Viruses in Pigs. In: Richt JA, Webby RJ, editors. Swine Influenza. Current Topics in Microbiology and Immunology. 370. Berlin: Springer-Verlag Berlin; 2013. p. 113-31.
56. Lorusso A, Vincent AL, Harland ML, Alt D, Bayles DO, Swenson SL, et al. Genetic and antigenic characterization of H1 influenza viruses from United States swine from 2008. Journal of General Virology. 2011;92:919-30.
57. Wolf YI, Viboud C, Holmes EC, Koonin EV, Lipman DJ. Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. Biology Direct. 2006;1.
58. Bahl J, Nelson MI, Chan KH, Chen R, Vijaykrishna D, Halpin RA, et al. Temporally structured metapopulation dynamics and persistence of influenza A H3N2

virus in humans. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(48):19359-64.
59. Worobey M, Han G-Z, Rambaut A. A synchronized global sweep of the internal genes of modern avian influenza virus. Nature. 2014;508(7495):254-+.
60. Furuse Y, Shimabukuro K, Odagiri T, Sawayama R, Okada T, Khandaker I, et al. Comparison of selection pressures on the HA gene of pandemic (2009) and seasonal human and swine influenza A H1 subtype viruses. Virology. 2010;405(2):314-21.
61. Nelson MI, Wentworth DE, Culhane MR, Vincent AL, Viboud C, LaPointe MP, et al. Introductions and Evolution of Human-Origin Seasonal Influenza A Viruses in Multinational Swine Populations. Journal of Virology. 2014;88(17):10110-9.
62. Guarnaccia T, Carolan LA, Maurer-Stroh S, Lee RTC, Job E, Reading PC, et al. Antigenic Drift of the Pandemic 2009 A(H1N1) Influenza Virus in a Ferret Model. Plos Pathogens. 2013;9(5).
63. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. Science. 2009;326(5953):734-6.
64. Romagosa A, Allerson M, Gramer M, Joo HS, Deen J, Detmer S, et al. Vaccination of influenza a virus decreases transmission rates in pigs. Veterinary Research. 2011;42.
65. Francis T. On the Doctrine of Original Antigenic Sin Proceedings of the American Philosophical Society: Proceedings of the; 1960. p. 572-8.
66. Gauger PC, Loving CL, Lager KM, Janke BH, Kehrli ME, Roth JA, et al. Vaccine-Associated Enhanced Respiratory Disease Does Not Interfere with the Adaptive Immune Response Following Challenge with Pandemic A/H1N1 2009. Viral Immunology. 2013;26(5):314-21.
67. East IJ, Todd PEE, Leach SJ. ORIGINAL ANTIGENIC SIN - EXPERIMENTS WITH A DEFINED ANTIGEN. Molecular Immunology. 1980;17(12):1539-44.
68. Zarnitsyna VI, Ellebedy AH, Davis C, Jacob J, Ahmed R, Antia R. Masking of antigenic epitopes by antibodies shapes the humoral immune response to influenza. Philos Trans R Soc Lond B Biol Sci. 2015;370(1676).
69. Kim JH, Skountzou I, Compans R, Jacob J. Original Antigenic Sin Responses to Influenza Viruses. Journal of Immunology. 2009;183(5):3294-301.
70. Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC. Influenza Virus Reassortment Occurs with High Frequency in the Absence of Segment Mismatch. Plos Pathogens. 2013;9(6).
71. Essere B, Yver M, Gavazzi C, Terrier O, Isel C, Fournier E, et al. Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(40):E3840-E8.
72. Lycett SJ, Baillie G, Coulter E, Bhatt S, Kellam P, McCauley JW, et al. Estimating reassortment rates in co-circulating Eurasian swine influenza viruses. Journal of General Virology. 2012;93:2326-36.

73. Murcia PR, Hughes J, Battista P, Lloyd L, Baillie GJ, Ramirez-Gonzalez RH, et al. Evolution of an Eurasian Avian-like Influenza Virus in Naive and Vaccinated Pigs. Plos Pathogens. 2012;8(5).
74. Wei K, Sun H, Sun Z, Sun Y, Kong W, Pu J, et al. Influenza A virus Acquires Enhanced Pathogenicity and Transmissibility After Serial Passages in Swine. J Virol. 2014.
75. Anderson TK, Nelson MI, Kitikoon P, Swenson SL, Korslund JA, Vincent AL. Population dynamics of cocirculating swine influenza A viruses in the United States from 2009 to 2012. Influenza and Other Respiratory Viruses. 2013;7:42-51.
76. Anderson TK, Campbell BA, Nelson MI, Lewis NS, Janas-Martindale A, Killian ML, et al. Characterization of co-circulating swine influenza A viruses in North America and the identification of a novel H1 genetic clade with antigenic significance. Virus Res. 2015;201:24-31.
77. Zhu HC, Webby R, Lam TTY, Smith DK, Peiris JSM, Guan Y. History of Swine Influenza Viruses in Asia. In: Richt JA, Webby RJ, editors. Swine Influenza. Current Topics in Microbiology and Immunology. 370. Berlin: Springer-Verlag Berlin; 2013. p. 57-68.
78. Poljak Z, Carman S, McEwen B. Assessment of seasonality of influenza in swine using field submissions to a diagnostic laboratory in Ontario between 2007 and 2012. Influenza and Other Respiratory Viruses. 2014;8(4):482-92.
79. Poljak Z, Dewey CE, Martin SW, Christensen J, Carman S, Friendship RM. Prevalence of and risk factors for influenza in southern Ontario swine herds in 2001 and 2003. Canadian Journal of Veterinary Research-Revue Canadienne De Recherche Veterinaire. 2008;72(1):7-17.
80. Choi YK, Goyal SM, Joo HS. Evaluation of transmission of swine influenza type A subtype H1N2 virus in seropositive pigs. American Journal of Veterinary Research. 2004;65(3):303-6.
81. Vincent AL, Ma W, Lager KM, Janke BH, Webby RJ, Garcia-Sastre A, et al. Efficacy of intranasal administration of a truncated NS1 modified live influenza virus vaccine in swine. Vaccine. 2007;25(47):7999-8009.
82. Kitikoon P, Gauger PC, Anderson TK, Culhane MR, Swenson S, Loving CL, et al. Swine influenza virus vaccine serologic cross-reactivity to contemporary US swine H3N2 and efficacy in pigs infected with an H3N2 similar to 2011-2012 H3N2v. Influenza and Other Respiratory Viruses. 2013;7:32-41.
83. Katz JM, Hancock K, Xu XY. Serologic assays for influenza surveillance, diagnosis and vaccine evaluation. Expert Review of Anti-Infective Therapy. 2011;9(6):669-83.
84. Slomka MJ, Densham ALE, Coward VJ, Essen S, Brookes SM, Irvine RM, et al. Real time reverse transcription (RRT)-polymerase chain reaction (PCR) methods for detection of pandemic (H1N1) 2009 influenza virus and European swine influenza A virus infections in pigs. Influenza and Other Respiratory Viruses. 2010;4(5):277-93.
85. Spackman E, Suarez DL. Type A influenza virus detection and quantitation by real-time RT-PCR. Methods Mol Biol. 2008;436:19-26.

86. Brown IH. History and Epidemiology of Swine Influenza in Europe. Swine Influenza. 2013;370:133-46.
87. Julkunen I, Pyhala R, Hovi T. ENZYME-IMMUNOASSAY, COMPLEMENT-FIXATION AND HEMAGGLUTINATION INHIBITION TESTS IN THE DIAGNOSIS OF INFLUENZA-A AND INFLUENZA-B VIRUS-INFECTIONS - PURIFIED HEMAGGLUTININ IN SUBTYPE-SPECIFIC DIAGNOSIS. Journal of Virological Methods. 1985;10(1):75-84.
88. Panyasing Y, Goodell CK, Gimenez-Lirola L, Kittawornrat A, Wang C, Schwartz KJ, et al. Kinetics of influenza A virus nucleoprotein antibody (IgM, IgA, and IgG) in serum and oral fluid specimens from pigs infected under experimental conditions. Vaccine. 2013;31(52):6210-5.
89. Ciacci-Zanella JR, Vincent AL, Prickett JR, Zimmerman SM, Zimmerman JJ. Detection of anti-influenza A nucleoprotein antibodies in pigs using a commercial influenza epitope-blocking enzyme-linked immunosorbent assay developed for avian species. Journal of Veterinary Diagnostic Investigation. 2010;22(1):3-9.
90. Direksin K, Joo H, Goyal SM. An immunoperoxidase monolayer assay for the detection of antibodies against swine influenza virus. Journal of Veterinary Diagnostic Investigation. 2002;14(2):169-71.
91. Pedersen JC. Hemagglutination-inhibition test for avian influenza virus subtype identification and the detection and quantitation of serum antibodies to the avian influenza virus. Methods Mol Biol. 2008;436:53-66.
92. Meguro H, Bryant JD, Torrence AE, Wright PF. CANINE KIDNEY CELL-LINE FOR ISOLATION OF RESPIRATORY VIRUSES. Journal of Clinical Microbiology. 1979;9(2):175-9.
93. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. Universal primer set for the full-length amplification of all influenza A viruses. Archives of Virology. 2001;146(12):2275-89.
94. Deng YM, Spirason N, Iannello P, Jelley L, Lau H, Barr IG. A simplified Sanger sequencing method for full genome sequencing of multiple subtypes of human influenza A viruses. Journal of Clinical Virology. 2015;68:43-8.
95. Lee HK, Tang JWT, Kong DHL, Koay ESC. Simplified Large-Scale Sanger Genome Sequencing for Influenza A/H3N2 Virus. Plos One. 2013;8(5).
96. Zhao JH, Grant SFA. Advances in Whole Genome Sequencing Technology. Current Pharmaceutical Biotechnology. 2011;12(2):293-305.
97. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
98. Archer J, Baillie G, Watson SJ, Kellam P, Rambaut A, Robertson DL. Analysis of high-depth sequence data for studying viral diversity: a comparison of next generation sequencing platforms using Segminator II. Bmc Bioinformatics. 2012;13.
99. Shao W, Boltz VF, Spindler JE, Kearney MF, Maldarelli F, Mellors JW, et al. Analysis of 454 sequencing error rate, error sources, and artifact recombination for detection of Low-frequency drug resistance mutations in HIV-1 DNA. Retrovirology. 2013;10.

100. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012;9(4):357-U54.
101. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology. 2009;10(3):10.
102. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
103. Zhou B, Donnelly ME, Scholes DT, George KS, Hatta M, Kawaoka Y, et al. Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. Journal of Virology. 2009;83(19):10309-13.
104. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008;92(5):255-64.
105. Van den Hoecke S, Verhelst J, Vuylsteke M, Saelens X. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. Bmc Genomics. 2015;16.
106. Bidzhieva B, Zagorodnyaya T, Karagiannis K, Simonyan V, Laassri M, Chumakov K. Deep sequencing approach for genetic stability evaluation of influenza A viruses. Journal of Virological Methods. 2014;199:68-75.
107. Schultz-Cherry S, Olsen CW, Easterday BC. History of Swine Influenza. In: Richt JA, Webby RJ, editors. Swine Influenza. Current Topics in Microbiology and Immunology. 370. Berlin: Springer-Verlag Berlin; 2013. p. 21-7.
108. Hinshaw VS, Bean WJ, Webster RG, Easterday BC. PREVALENCE OF INFLUENZA-VIRUSES IN SWINE AND ANTIGENIC AND GENETIC RELATEDNESS OF INFLUENZA-VIRUSES FROM MAN AND SWINE. Virology. 1978;84(1):51-62.
109. Chambers TM, Hinshaw VS, Kawaoka Y, Easterday BC, Webster RG. INFLUENZA VIRAL-INFECTION OF SWINE IN THE UNITED-STATES 1988-1989. Archives of Virology. 1991;116(1-4):261-5.
110. Olsen CW, Carey S, Hinshaw L, Karasin AI. Virologic and serologic surveillance for human, swine and avian influenza virus infections among pigs in the north-central United States. Archives of Virology. 2000;145(7):1399-419.
111. Olsen CW. The emergence of novel swine influenza viruses in North America. Virus Research. 2002;85(2):199-210.
112. Webby RJ, Rossow K, Erickson G, Sims Y, Webster R. Multiple lineages of antigenically and genetically diverse influenza A virus co-circulate in the United States swine population. Virus Research. 2004;103(1-2):67-73.
113. Corzo CA, Gramer M, Kuhn M, Mohr M, Morrison R. Observations regarding influenza A virus shedding in a swine breeding farm after mass vaccination. Journal of Swine Health and Production. 2012;20(6):283-9.
114. Meiners C, Loesken S, Doehring S, Starick E, Pesch S, Maas A, et al. Field study on swine influenza virus (SIV) infection in weaner pigs and sows. Tierarztl Prax Ausg G Grosstiere Nutztiere. 2014;42(6).

115. Sarkar D. Lattice Multivariate Data Visualization with R Introduction. Lattice: Multivariate Data Visualization with R. 2008.
116. Lesnoff M, Lancelot R. aod: Analysis of Over Dispersed Data. R package version 1.3. 2012.
117. Wickham H. ggplot2: elegant graphics for data analysis.: Springer New York; 2009.
118. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. 2014.
119. Revelle W. psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston Illinois, USA2014.
120. Van Reeth K, Brown IH, Durrwald R, Foni E, Labarque G, Lenihan P, et al. Seroprevalence of H1N1, H3N2 and H1N2 influenza viruses in pigs in seven European countries in 2002-2003. Influenza and Other Respiratory Viruses. 2008;2(3):99-105.
121. Maldonado J, Van Reeth K, Riera P, Sitja M, Saubi N, Espuna E, et al. Evidence of the concurrent circulation of H1N2, H1N1 and H3N2 influenza A viruses in densely populated pig areas in Spain. Veterinary Journal. 2006;172(2):377-81.
122. Kyriakis CS, Rose N, Foni E, Maldonado J, Loeffen WLA, Madec F, et al. Influenza A virus infection dynamics in swine farms in Belgium, France, Italy and Spain, 2006-2008. Veterinary Microbiology. 2013;162(2-4):543-50.
123. Dugan VG, Chen R, Spiro DJ, Sengamalay N, Zaborsky J, Ghedin E, et al. The evolutionary genetics and emergence of avian influenza viruses in wild birds. Plos Pathogens. 2008;4(5).
124. Alexander DJ, Capua I. Avian influenza in poultry. Worlds Poultry Science Journal. 2008;64(4):513-31.
125. Corzo CA, Gramer M, Lauer D, Davies PR. Prevalence and Risk Factors for H1N1 and H3N2 Influenza A Virus Infections in Minnesota Turkey Premises. Avian Diseases. 2012;56(3):488-93.
126. Murcia PR, Baillie GJ, Daly J, Elton D, Jervis C, Mumford JA, et al. Intra- and interhost evolutionary dynamics of equine influenza virus. J Virol. 84. United States2010. p. 6943-54.
127. USDA. United States Department of Agriculture (USDA). Economic Research Service (ERS). Hogs and Pork: Background 2014 [Available from: http://www.ers.usda.gov/topics/animal-products/hogs-pork/background.aspx.
128. Diaz A, Perez A, Sreevatsan S, Davies P, Culhane M, Torremorell M. Association between Influenza A Virus Infection and Pigs Subpopulations in Endemically Infected Breeding Herds. PLoS One. 2015;10(6):e0129213.
129. Babraham_Bioinformatics. FastQC. http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/2014.
130. Langdon WB. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. Biodata Mining. 2015;8.
131. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Tatusova T. FLAN: a web server for influenza virus genome annotation. Nucleic Acids Res. 2007;35(Web Server issue):W280-4.

132. Squires RB, Noronha J, Hunt V, Garcia-Sastre A, Macken C, Baumgarth N, et al. Influenza Research Database: an integrated bioinformatics resource for influenza research and surveillance. Influenza and Other Respiratory Viruses. 2012;6(6):404-16.
133. Zell R, Scholtissek C, Ludwig S. Genetics, Evolution, and the Zoonotic Capacity of European Swine Influenza Viruses. In: Richt JA, Webby RJ, editors. Swine Influenza. Current Topics in Microbiology and Immunology. 3702013. p. 29-55.
134. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. Bmc Bioinformatics. 2004;5:1-19.
135. Price MN, Dehal PS, Arkin AP. FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Plos One. 2010;5(3).
136. Yang ZH. MAXIMUM-LIKELIHOOD PHYLOGENETIC ESTIMATION FROM DNA-SEQUENCES WITH VARIABLE RATES OVER SITES - APPROXIMATE METHODS. Journal of Molecular Evolution. 1994;39(3):306-14.
137. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and clustal X version 2.0. Bioinformatics. 2007;23(21):2947-8.
138. Gleed ML, Busath DD. Why bound amantadine fails to inhibit proton conductance according to simulations of the drug-resistant influenza A M2 (S31N). J Phys Chem B. 2015;119(3):1225-31.
139. Vincent AL, Lager KM, Ma WJ, Lekcharoensuk P, Gramer MR, Loiacono C, et al. Evaluation of hemagglutinin subtype 1 swine influenza viruses from the United States. Veterinary Microbiology. 2006;118(3-4):212-22.
140. Diaz A, Allerson M, Culhane M, Sreevatsan S, Torremorell M. Antigenic drift of H1N1 influenza A virus in pigs with and without passive immunity. Influenza and Other Respiratory Viruses. 2013;7:52-60.
141. Diaz A, Enomoto S, Romagosa A, Sreevatsan S, Nelson M, Culhane M, et al. Genome plasticity of triple reassortant H1N1 influenza A virus during infection of vaccinated pigs. J Gen Virol. 2015.
142. Kitikoon P, Strait EL, Thacker EL. The antibody responses to swine influenza virus (SIV) recombinant matrix 1 (rM1), matrix 2 (M2), and hemagglutinin (HA) proteins in pigs with different SIV exposure. Veterinary Microbiology. 2008;126(1-3):51-62.
143. Kitikoon P, Nilubol D, Erickson BJ, Janke BH, Hoover TC, Sornsen SA, et al. The immune response and maternal antibody interference to a heterologous H1N1 swine influenza virus infection following vaccination. Veterinary Immunology and Immunopathology. 2006;112(3-4):117-28.
144. Choi MJ, Torremorell M, Bender JB, Smith K, Boxrud D, Ertl JR, et al. Live animal markets in Minnesota: a potential source for emergence of novel influenza A viruses and interspecies transmission. Clin Infect Dis. 2015.
145. Baranovich T, Bahl J, Marathe BM, Culhane M, Stigger-Rosser E, Darnell D, et al. Influenza A viruses of swine circulating in the United States during 2009-2014 are susceptible to neuraminidase inhibitors but show lineage-dependent resistance to adamantanes. Antiviral Res. 2015;117:10-9.
146. Krumbholz A, Schmidtke M, Bergmann S, Motzke S, Bauer K, Stech J, et al. High prevalence of amantadine resistance among circulating European porcine influenza A viruses. J Gen Virol. 2009;90(Pt 4):900-8.

147. Bright RA, Medina MJ, Xu X, Perez-Oronoz G, Wallis TR, Davis XM, et al. Incidence of adamantane resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern. Lancet. 2005;366(9492):1175-81.
148. Deyde VM, Xu X, Bright RA, Shaw M, Smith CB, Zhang Y, et al. Surveillance of resistance to adamantanes among influenza A(H3N2) and A(H1N1) viruses isolated worldwide. J Infect Dis. 2007;196(2):249-57.
149. Drummond A, Pybus OG, Rambaut A. Inference of viral evolutionary rates from molecular sequences. Advances in Parasitology, Vol 54. 2003;54:331-58.
150. Fanning TG, Reid AH, Taubenberger JK. Influenza A virus neuraminidase: Regions of the protein potentially involved in virus-host interactions. Virology. 2000;276(2):417-23.
151. Long JX, Bushnell RV, Tobin JK, Pan KY, Deem MW, Nara PL, et al. Evolution of H3N2 Influenza Virus in a Guinea Pig Model. Plos One. 2011;6(7).
152. Van Reeth K. Avian and swine influenza viruses: our current understanding of the zoonotic risk. Veterinary Research. 2007;38(2):243-60.
153. Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution. 1999;16(1):37-48.
154. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006;22(2):195-201.
155. Deem MW, Pan KY. The epitope regions of H1-subtype influenza A, with application to vaccine efficacy. Protein Engineering Design & Selection. 2009;22(9):543-6.
156. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. THE ANTIGENIC STRUCTURE OF THE INFLUENZA-VIRUS A/PR/8/34 HEMAGGLUTININ (H-1 SUBTYPE). Cell. 1982;31(2):417-27.
157. Kuroda M, Katano H, Nakajima N, Tobiume M, Ainai A, Sekizuka T, et al. Characterization of Quasispecies of Pandemic 2009 Influenza A Virus (A/H1N1/2009) by De Novo Sequencing Using a Next-Generation DNA Sequencer. Plos One. 2010;5(4).
158. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, et al. The global circulation of seasonal influenza A (H3N2) viruses. Science. 2008;320(5874):340-6.
159. Blanc A, Ruchansky D, Clara M, Achaval F, Le Bas A, Arbiza J. Serologic Evidence of Influenza A and B Viruses in South American Fur Seals (Arctocephalus australis). Journal of Wildlife Diseases. 2009;45(2):519-21.
160. Simonsen L, Spreeuwenberg P, Lustig R, Taylor RJ, Fleming DM, Kroneman M, et al. Global mortality estimates for the 2009 Influenza Pandemic from the GLaMOR project: a modeling study. PLoS Med. 2013;10(11):e1001558.
161. Domingo E, Sheldon J, Perales C. Viral Quasispecies Evolution. Microbiology and Molecular Biology Reviews. 2012;76(2):159-216.
162. Lemey P, Rambaut A, Pybus OG. HIV evolutionary dynamics within and among hosts. Aids Reviews. 2006;8(3):125-40.
163. Belshaw R, Pybus OG, Rambaut A. The evolution of genome compression and genomic novelty in RNA viruses. Genome Research. 2007;17(10):1496-504.

164. de Groot J, Ruis MA, Scholten JW, Koolhaas JM, Boersma WJ. Long-term effects of social stress on antiviral immunity in pigs. Physiol Behav. 73. United States2001. p. 145-58.
165. Domingo E, Martinezsalas E, Sobrino F, Delatorre JC, Portela A, Ortin J, et al. THE QUASISPECIES (EXTREMELY HETEROGENEOUS) NATURE OF VIRAL-RNA GENOME POPULATIONS - BIOLOGICAL RELEVANCE - A REVIEW. Gene. 1985;40(1):1-8.
166. Salemi M. The intra-host evolutionary and population dynamics of human immunodeficiency virus type 1: a phylogenetic perspective. Infect Dis Rep. 2013;5(Suppl 1):e3.
167. Tu Z, He YL, Lu H, Xu L, Yang ZB, Yang C, et al. Mutant spectrum of dengue type 1 virus in the plasma of patients from the 2006 epidemic in South China. Int J Infect Dis. 2013;17(11):e1080-1.
168. Debbink K, Lindesmith LC, Ferris MT, Swanstrom J, Beltramello M, Corti D, et al. Within-host evolution results in antigenically distinct GII.4 noroviruses. J Virol. 88. United States: American Society for Microbiology. All Rights Reserved.; 2014. p. 7244-55.
169. Bourret V, Croville G, Mansuy JM, Mengelle C, Mariette J, Klopp C, et al. Intra-host viral variability in children clinically infected with H1N1 (2009) pandemic influenza. Infect Genet Evol. 2015;33:47-54.
170. Tao H, Steel J, Lowen AC. Intrahost dynamics of influenza virus reassortment. J Virol. 88. United States: American Society for Microbiology. All Rights Reserved.; 2014. p. 7485-92.
171. Roedig JV, Rapp E, Hoper D, Genzel Y, Reichl U. Impact of Host Cell Line Adaptation on Quasispecies Composition and Glycosylation of Influenza A Virus Hemagglutinin. Plos One. 2011;6(12).
172. Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, et al. Characterizing and measuring bias in sequence data. Genome Biology. 2013;14(5).
173. Cummings SM, McMullan M, Joyce DA, van Oosterhout C. Solutions for PCR, cloning and sequencing errors in population genetic analysis. Conservation Genetics. 2010;11(3):1095-7.
174. Chen-Harris H, Borucki MK, Torres C, Slezak TR, Allen JE. Ultra-deep mutant spectrum profiling: improving sequencing accuracy using overlapping read pairs. Bmc Genomics. 2013;14.
175. Bourret V, Croville G, Mariette J, Klopp C, Bouchez O, Tiley L, et al. Whole-genome, deep pyrosequencing analysis of a duck influenza A virus evolution in swine cells. Infection Genetics and Evolution. 2013;18:31-41.
176. Lundberg KS, Shoemaker DD, Adams MWW, Short JM, Sorge JA, Mathur EJ. HIGH-FIDELITY AMPLIFICATION USING A THERMOSTABLE DNA-POLYMERASE ISOLATED FROM PYROCOCCUS-FURIOSUS. Gene. 1991;108(1):1-6.
177. Andre P, Kim A, Khrapko K, Thilly WG. Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. Genome Research. 1997;7(8):843-52.

178. Hedman J, Nordgaard A, Rasmusson B, Ansell R, Radstrom P. Improved forensic DNA analysis through the use of alternative DNA polymerases and statistical modeling of DNA profiles. Biotechniques. 2009;47(5):951-8.
179. Wu JY, Jiang XT, Jiang YX, Lu SY, Zou F, Zhou HW. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. Bmc Microbiology. 2010;10.
180. Ramakrishnan MA, Tu ZJ, Singh S, Chockalingam AK, Gramer MR, Wang P, et al. The feasibility of using high resolution genome sequencing of influenza A viruses to detect mixed infections and quasispecies. PLoS One. 2009;4(9):e7105.
181. Goto N, Prins P, Nakao M, Bonnal R, Aerts J, Katayama T. BioRuby: bioinformatics software for the Ruby programming language. Bioinformatics. 2010;26(20):2617-9.
182. Korber B. HIV Signature and Sequence Variation Analysis in: Computational Analysis of HIV Molecular Sequences. Netherlands: Kluwer Academic Publishers.; 2000.
183. Nei M, Gojobori T. SIMPLE METHODS FOR ESTIMATING THE NUMBERS OF SYNONYMOUS AND NONSYNONYMOUS NUCLEOTIDE SUBSTITUTIONS. Molecular Biology and Evolution. 1986;3(5):418-26.
184. Short KR, Kroeze E, Fouchier RAM, Kuiken T. Pathogenesis of influenza-induced acute respiratory distress syndrome. Lancet Infectious Diseases. 2014;14(1):57-69.
185. Davidson I, Al-Touri A, Raibshtein I, Hadani Y, Bombarov V, Yadin H, et al. Seroepidemiology Survey and Isolation of Swine Influenza Viruses from Subclinical Infections in Israel During the Years 2009-2011. Israel Journal of Veterinary Medicine. 2014;69(2):62-7.
186. Meurens F, Summerfield A, Nauwynck H, Saif L, Gerdts V. The pig: a model for human infectious diseases. Trends in Microbiology. 2012;20(1):50-7.
187. Simon-Grife M, Martin-Valls GE, Vilar MJ, Busquets N, Mora-Salvatierra M, Bestebroer TM, et al. Swine influenza virus infection dynamics in two pig farms; results of a longitudinal assessment. Veterinary Research. 2012;43.
188. Domingo E, Holland JJ. RNA virus mutations and fitness for survival. Annual Review of Microbiology. 1997;51:151-78.
189. Handel A, Rohani P. Crossing the scale from within-host infection dynamics to between-host transmission fitness: a discussion of current assumptions and knowledge. Philosophical Transactions of the Royal Society B-Biological Sciences. 2015;370(1675):10.
190. Camacho A, Ballesteros S, Graham AL, Carrat F, Ratmann O, Cazelles B. Explaining rapid reinfections in multiple-wave influenza outbreaks: Tristan da Cunha 1971 epidemic as a case study. Proceedings of the Royal Society B-Biological Sciences. 2011;278(1725):3635-43.
191. Nakajima S, Nishikawa F, Nakamura K, Nakajima K. Analysis of influenza A virus reinfection in children in Japan during 1983-91. Epidemiology and Infection. 1995;115(3):591-601.

192. Frank AL, Taber LH, Glezen WP, Paredes A, Couch RB. RE-INFECTION WITH INFLUENZA-A (H3N2) VIRUS IN YOUNG-CHILDREN AND THEIR FAMILIES. Journal of Infectious Diseases. 1979;140(6):829-36.
193. Pepin KM, Volkov I, Banavar JR, Wilke CO, Grenfell BT. Phenotypic differences in viral immune escape explained by linking within-host dynamics to host-population immunity. Journal of Theoretical Biology. 2010;265(4):501-10.
194. Romagosa A, Gramer M, Joo HS, Torremorell M. Sensitivity of oral fluids for detecting influenza A virus in populations of vaccinated and non-vaccinated pigs. Influenza and Other Respiratory Viruses. 2012;6(2):110-8.
195. Loeffen WL, Heinen PP, Bianchi AT, Hunneman WA, Verheijden JH. Effect of maternally derived antibodies on the clinical signs and immune response in pigs after primary and secondary infection with an influenza H1N1 virus. Vet Immunol Immunopathol. 2003;92(1-2):23-35.
196. Pan KY. Understanding Original Antigenic Sin in Influenza with a Dynamical System. Plos One. 2011;6(8).
197. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GCM, Vervaet G, et al. Substitutions Near the Receptor Binding Site Determine Major Antigenic Change During Influenza Virus Evolution. Science. 2013;342(6161):976-9.
198. Li Y, Bostick DL, Sullivan CB, Myers JL, Griesemer SB, George KS, et al. Single Hemagglutinin Mutations That Alter both Antigenicity and Receptor Binding Avidity Influence Influenza Virus Antigenic Clustering. Journal of Virology. 2013;87(17):9904-10.
199. Lewis NS, Anderson TK, Kitikoon P, Skepner E, Burke DF, Vincent AL. Substitutions near the Hemagglutinin Receptor-Binding Site Determine the Antigenic Evolution of Influenza A H3N2 Viruses in US Swine. Journal of Virology. 2014;88(9):4752-63.
200. Gabriel G, Arck PC. Sex, Immunity and Influenza. Journal of Infectious Diseases. 2014;209:S93-S9.
201. Klein SL, Huber S. Sex Differences in Susceptibility to Viral Infection. In: Klein SL, Roberts CW, editors. Sex Hormones and Immunity to Infection. Berlin: Springer-Verlag Berlin; 2010. p. 93-122.
202. Jensen-Fangel S, Mohey R, Johnsen SP, Andersen PL, Sorensen HT, Ostergaard L. Gender differences in hospitalization rates for respiratory tract infections in Danish youth. Scandinavian Journal of Infectious Diseases. 2004;36(1):31-6.
203. Gagnon A, Acosta JE, Madrenas J, Miller MS. Is Antigenic Sin Always "Original?" Re-examining the Evidence Regarding Circulation of a Human H1 Influenza Virus Immediately Prior to the 1918 Spanish Flu. Plos Pathogens. 2015;11(3):6.
204. Vincent AL, Ma W, Lager KM, Richt JA, Janke BH, Sandbulte MR, et al. Live attenuated influenza vaccine provides superior protection from heterologous infection in pigs with maternal antibodies without inducing vaccine-associated enhanced respiratory disease. J Virol. 2012;86(19):10597-605.